BEYOND THE SURFACE: ENHANCING FRESHWATER POND ECOSYSTEM ASSESSMENT THROUGH eDNA METABARCODING

Claire E. Robertson

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (BIOSCIENCES)

MARCH 2024

Summary

Freshwater ponds are small, lentic freshwater habitats. Composing an estimated 30 % of global standing water area, they are important reservoirs of freshwater biodiversity providing key ecosystem services. Environmental DNA (eDNA) metabarcoding is a novel method of sampling biodiversity, where an organism's presence is inferred by sampling and identifying their DNA present in the environment. This method has advantages in terms of efficiency, scalability and identifying small and cryptic species.

This PhD examines the use of eDNA metabarcoding to monitor freshwater ponds. Firstly, I developed new techniques, including using genetic markers never previously used in ponds or other freshwater environments (Chapter 2). This increased the taxonomic breadth of pond monitoring by targeting prokaryotic and eukaryotic microbes, fungi, microfauna and algae alongside the traditionally used macrophytes and macroinvertebrates.

I examined seasonal dynamics of communities in pond water and sediment via eDNA metabarcoding (Chapter 3). I also used eDNA metabarcoding to test ecological hypotheses at the landscape level (Chapter 4), investigating environmental and spatial drivers of pond community composition. I compared community metrics derived from metabarcoding data with a traditional pond ecosystem assessment tool (the Habitat Suitability Index for Great Crested Newts).

Pond communities in water and sediment were significantly dissimilar, and this difference was bigger for microbes than for larger organisms. Communities in pond water displayed significant seasonal turnover, but this was less pronounced in sediment communities. Higher plant eDNA sequence reads from water samples were over 70 % terrestrial in origin. Prokaryotic communities were structured by pond water chemistry and surrounding land use, and their composition differed in ponds of different HSI levels. Green plant and algal communities were spatially structured, and other eukaryotes showed no clear structuring pattern.

These findings encourage a shift in eDNA surveying of ponds, demonstrating that bacteria communities in sediments are good bioindicators of overall pond ecosystem health.

Contents

Summary i
Acknowledgements iv
Introduction 1
Literature Review: Molecular insights into pond ecosystems: current applications and future prospects 4
Abstract 4
Section 1: Pond ecology, status and conservation 4
Section 2: eDNA metabarcoding in freshwater ponds 16
Section 3: Bioindicators in freshwaters, ponds and using eDNA 28
Section 4: Community assembly and eDNA 35
Conclusion 45
Chapter 2: From microbes to macrophytes: optimising multi-taxa eDNA metabarcoding methods in freshwater ponds 48
Abstract 48
Introduction 49
Methods 53
Results 63
Discussion 76
Conclusions 81
Chapter 3: Understanding seasonal turnover in pond ecosystems using eDNA metabarcoding 84
Abstract 84
Introduction 85
Hypotheses 88
Methods 89
Results 97
Discussion 113
Conclusion 117
Chapter 4: Evaluating environmental and spatial drivers of community composition in lowland ponds: insights from eDNA metabarcoding. 119
Abstract 119
Introduction 120
Methods 125
Results 137
Discussion 154

__Conclusions 157
Chapter 5: Conclusions 159
__Limitations 159
__Methodological findings 160
__Taxonomic and bioindicator findings 164
__Ecological findings 165
__Conclusions 167
References 169

Acknowledgements

Thanks go firstly to Dan Read for granting me this opportunity and for all the ideas, nudges in the right direction, proof-reading and for being a supportive, and undemanding, supervisor. Profuse thanks to Tim Goodall, without whom the lab work, and therefore the entire PhD, would have been totally impossible! Thanks to others in the Molecular Ecology group who helped me out when I was stuck.

Next, thank you to all my volunteer field samplers who helped with great enthusiasm, and often at a 2m distance: Amy, Cara, Nikki, Anna, Clemency, Katherine and especially Nick. I really enjoyed traipsing around the countryside, spotting dragonflies, stonewort and snipe and wading in deep sulphurous silt with you.

Thanks to Jeremy Biggs, Penny Williams and others at the Freshwater Habitats Trust for letting me muck about in what I think of as “your” ponds. Thank you, Steve Ormerod, Isabelle Durance and the entire FRESH DTP team for amazing training and opportunities and opening my eyes to the wonders of freshwater.

Thank you to my friends and family, especially George, for believing I could do it when I wanted to give up.

I’d like to dedicate this to the memory of my Nan, Joan Robertson, and my Dad, Andy Robertson. For their encouragement and love, and for teaching me to wonder at the natural world.

This research was supported by funding from the National Environment Research Council and the UK Centre for Ecology and Hydrology.

Introduction

Freshwater ponds are small, lentic freshwater habitats less than five metres in depth and five hectares in area (Richardson et al. 2022). Historically, ponds were not a primary focus of freshwater science and monitoring. Larger rivers, lakes, and wetlands took precedence, as they were presumed to play a more significant role in global biodiversity, carbon and nutrient cycles, water dynamics, and ecosystem services (Hill et al. 2018).

The past decade has seen a surge of interest in ponds. Improvements in satellite imagery have allowed better estimates of pond numbers and area (around 30 % of global standing water area, Verpoorter et al. 2014, Holgerson & Raymond 2016), and a clearer definition of ponds as distinct from other freshwater habitats, based on ecosystem functioning (Richardson et al. 2022).

In many areas, ponds are crucial for freshwater biodiversity, containing more species at a landscape level than rivers or lakes (Davies et al. 2008, Richardson et al. 2014, Bolgovics et al. 2019). The relevance of ponds for local, regional, and global methane emissions (Rosentreter et al. 2021), pollutant and nutrient reduction (Tournebize et al. 2016) and water storage and cycling (Golden et al. 2015) is actively being investigated, and ponds have been put forward as key nature-based solutions to biodiversity loss, flooding and drought and climate change (Cuenca-Cambronero et al. 2023).

In light of the renewed recognition of the significance of these small habitats, the UK is currently working on developing new monitoring schemes for ponds (Natural England 2019, DEFRA 2022). This initiative aims to address previous shortcomings in the inconsistent and sporadic monitoring of ponds and the omission of ponds from regulatory frameworks for freshwater protection (Hill et al. 2018).

Environmental DNA (eDNA) metabarcoding is a novel method of sampling biodiversity. Instead of visually counting or trapping individual organisms, as in traditional surveying methods, an organism’s presence is inferred by sampling and identifying their DNA which is present in the environment (Tablerlet et al. 2018), which may be water, soil, sediment or even air. In the past decade, the number of eDNA metabarcoding studies has rapidly

increased, aided by technological developments in next-generation sequencing which make it possible to determine the sequences of thousands of eDNA copies simultaneously.

eDNA techniques are being progressively employed in regulatory biodiversity monitoring, e.g., for protected and invasive species, and as part of citizen science schemes (see Chapter 1, section 3). Over the past four years (the length of this PhD study), there have been significant developments in eDNA applications such as passive sampling (Bessey et al. 2021), taxonomy-free metrics (Cordier et al. 2020) and machine-learning (Fruhe et al. 2021). eDNA metabarcoding is continually being applied in new environments and across various regions of the world.

This PhD examines the use of eDNA metabarcoding for monitoring freshwater ponds in lowland England. During this PhD, I developed new techniques for monitoring ponds using eDNA metabarcoding, including using genetic markers never previously used in ponds or other freshwater environments. I increased the taxonomic breadth of pond monitoring by targeting bacteria, fungi, microfauna and algae alongside the macrophytes and macroinvertebrates traditionally used for ecosystem assessments of these environments. I examined the seasonal dynamics of communities in pond water and sediment using eDNA metabarcoding, expanding the temporal resolution of traditional pond surveying. Finally, I used eDNA metabarcoding to test ecological hypotheses at the landscape level.

My two key study sites were Pinkhill Meadows, a five-hectare pond complex and nature reserve in Oxfordshire, and 31 ponds across a roughly 500 km^2 area of lowland central England (Oxfordshire, Buckinghamshire and Northamptonshire). Fieldwork for the Pinkhill Meadow studies (Chapters 2 and 3) was carried out from January to November 2020 (inclusive). Fieldwork for the landscape study (Chapter 4) was carried out in June and July 2022.

This PhD opens with a literature review (Chapter 1), covering the ponds in the UK and internationally (their ecology, numbers, status and threats), the status of eDNA metabarcoding in freshwater environments, biodiversity monitoring and bioindicators and freshwater community assembly.

In Chapter 2, I report on the results of an eDNA metabarcoding study using five different genetic markers across 20 ponds in Pinkhill Meadow, sampled on a single day in June 2020.

As well as detailing my taxonomic results, I also assess the benefits and drawbacks of various methods in field, in the laboratory and *in silico,* by comparing results from different genetic markers, from merged and separate water samples, and from using different bioinformatic processing techniques.

Chapter 3 examines the seasonal dynamics of pond communities in the same 20 ponds but over five different days in 2020 (in January, March, July, September and November), via eDNA metabarcoding using three different genetic markers. I assess the differences in seasonal dynamics between microbes, multicellular organisms and plants, and between communities in pond water and sediment.

Chapter 4 reports on results from a landscape-scale eDNA metabarcoding study of 31 ponds. Again, three different genetic markers were used to target prokaryotic microbes, eukaryotes and green plants and algae. The impact of pond physio-chemistry, surrounding landscape and spatial effects on these three communities is assessed using variation partitioning and constrained ordination. The potential for eDNA metabarcoding to provide bioindicators of pond ecological quality is also explored, through indicator species analysis and comparing community composition with a traditional pond health metric, the Habitat Suitability Index for Great Crested Newts.

Finally, I conclude by outlining the key findings of my work in the Conclusions chapter. I focus on my contribution to expanding knowledge of the taxonomic composition of pond environments, implications for biomonitoring of ponds using eDNA metabarcoding, and environmental and spatial drivers of pond community assembly.

Literature Review: Molecular insights into pond ecosystems: current applications and future prospects

Abstract

Global biodiversity is experiencing steep and alarming declines in species richness and abundance. Freshwater ecosystems and species are amongst the most threatened. Ponds (freshwater lentic habitats <5 ha and <5 m deep) and other smaller freshwater habitats are understudied and less well protected compared to other components of the freshwater environment, yet available evidence suggests their contribution to freshwater biodiversity outweighs their small size. The first section of this literature review briefly describes the ecology, distribution, status and threats to ponds in the UK and internationally. The second describes molecular ecological approaches to studying ponds, and their current and future applications. Next, bioindicators of freshwater ecological quality are examined, and future research directions in both ponds and molecular ecology in general are recommended. The final section examines the literature surrounding freshwater community assembly and the importance of freshwater connectivity to biodiversity.

Section 1: Pond ecology, status and conservation

Introduction

It is now widely accepted that we are in the midst of an anthropogenically-driven global biodiversity crisis: species extinctions are 100-1000 the background rate experienced over much of the Earth's history, leading to assertions we are experiencing a "sixth mass extinction" of biological diversity (Chapin et al. 2000, Mace et al. 2005, Rockstrom et al. 2009). A recent report showed that one million species – 11.5 % of the estimated global total

- are threatened with extinction, but this proportion varies for different groups e.g., an average of 25 % for well-studied groups (IPBES, 2019).

Freshwater ecosystems are among the most threatened ecosystems worldwide. Despite lakes, reservoirs and rivers occupying 2.3 % of the Earth's surface, freshwaters support an estimated 9.5 % of all described animal species (Reid et al. 2019). However, population declines in freshwater far outpace those in terrestrial and marine ecosystems: the Living Planet Index estimates a decline of 83 % since 1970, compared to 36 % and 38 % for land and sea respectively (WWF, 2018). Ponds conform to this wider pattern: despite constituting a relatively small proportion of total freshwater habitat worldwide, ponds support a high level of freshwater biodiversity (Biggs et al. 2017). They are also in decline (Smith et al. 2022) and many species they harbour are endangered (Hill et al. 2018).

What is a pond?

Many authors have struggled to define precisely what a pond is. Past attempts have focused on what features of ponds differentiate them from lakes, for instance a lack of thermal stratification (e.g., Macan and Worthington 1972), or shallow enough for light to penetrate to the bottom throughout (e.g., Fitter and Manuel 1986), although exceptions are found for both. There is no clear size boundary between ponds and puddles at the lower bound, and lakes at the upper bound: the Ramsar convention defines anything <8 ha as a pond (Ramsar 2022), yet the state of Wisconsin considers waterbodies <0.1 ha lakes (Wisconsin DNR 2009). The issue is muddled by the many English words for small, lentic freshwater habitats (e.g., prairie potholes, impoundments, small wetlands, vernal pools etc.)

The lack of a clear definition has hampered research, and in many studies habitats which fit the above criteria are referred to as "small wetlands" or "shallow lakes", or ponds are lumped together with lakes with an assumption that patterns of biodiversity and ecosystem function will be the same for both habitats (e.g., Macingo et al. 2019). However, a recent paper found that many metrics of freshwater ecosystem function, such as water chemistry and gas fluxes, changed in a non-linear fashion with waterbody size, depth, and percentage cover of emergent vegetation. The authors put forward a definition of a pond as "a lentic waterbody which is less than five hectares in area, less than five meters deep and with less than 30 % emergent vegetation cover" (Richardson et al. 2022), with lakes being larger and deeper and wetlands having a greater proportion of vegetation cover.

Ponds may be created naturally or manmade. It's estimated that at the end of the last ice age, approximately 12,000 years before present, 25 % of Great Britain was a wetland of

some kind, and the density of water bodies was as much as 100-200 per km ^{2}. Now it is around 2 per km ^{2}, and most are manmade (Biggs & Williams 2024, see Fig 1). The Lowland Pond Survey in 1996 estimated only 2 % of ponds were formed by natural processes (Williams et al. 1998b).

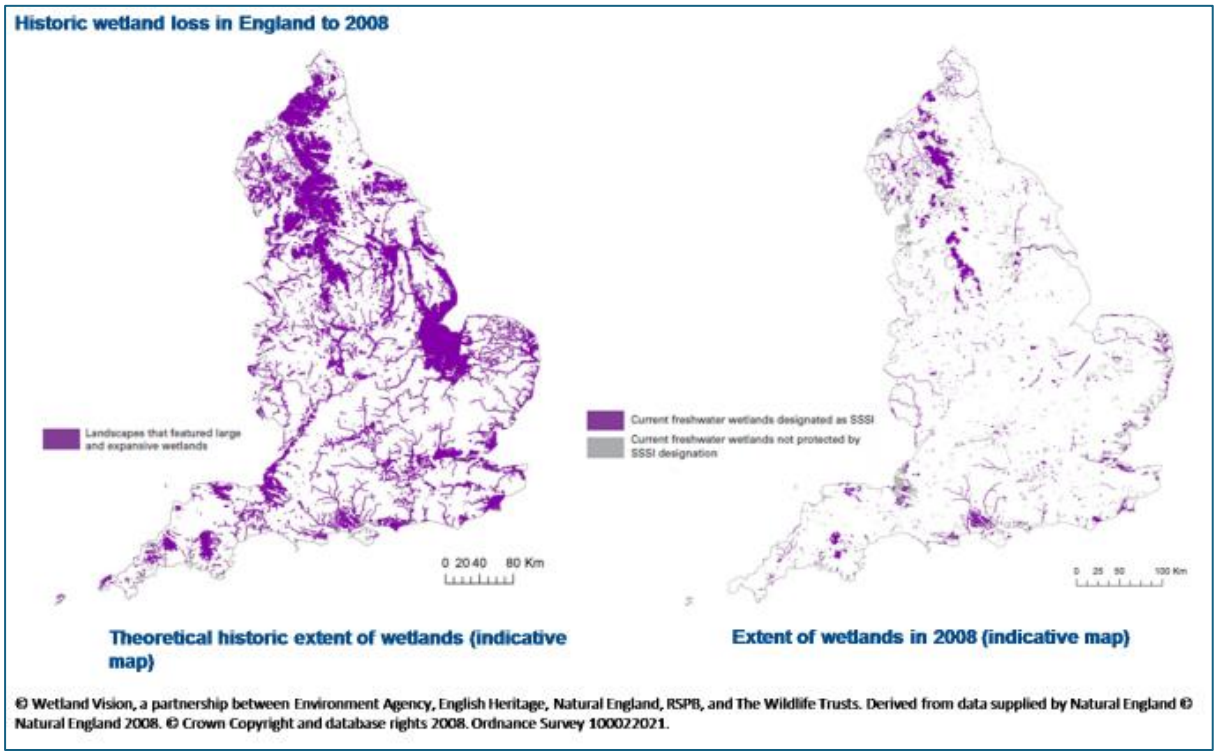


Figure 1.1: Historical extent of wetlands in England compared to extent of wetlands in 2008. From RSPB (2008) "A 50 Year Vision for Wetlands".

Natural pond creation in the past varied widely, occurring in natural hollows, dune slacks, beaver dams, tree falls, and river floodplains with meandering rivers leaving abandoned channels (Biggs & Williams, 2023 pp56-57). However, extensive land drainage, ploughing, tree felling, and channel modification have almost halted natural pond formation in Britain. Over centuries, ponds have been intentionally created by humans for diverse purposes, such as soil improvement (marling), water supply for various needs, aesthetic enjoyment (e.g., moats), industrial power (e.g., mill ponds), and as a by-product of mineral extraction (e.g., gravel ponds) (Biggs & Williams, 2023 pp78-96).

Research indicates no fundamental community differences between natural and manmade ponds. For example, the National Pond Survey in 1998 demonstrated that pond communities

could mainly be categorized by water acidity, with geological, depth, and permanence considerations outweighing the natural or manmade origin of the pond (Biggs et al., 2000; Biggs & Williams, 2023 p99). However, there is conflicting evidence: studies in southern Spain found significant differences in alpha and gamma diversity for zooplankton and macroinvertebrates between ponds with an artificial substratum (concrete or plastic) and ponds with a natural substratum (Fuentes-Rodrigues et al. 2013, Leon et al. 2008). Hill et al. (2017) compared hundreds of urban and non-urban ponds and found marked differences in community composition. Garden ponds also have lower biodiversity value than their 'wild' counterparts, due to their small size, frequent macrophyte removal and presence of fish and water features (Hill et al. 2021)

Ponds are generally considered successional habitats, evolving from open water to swamp, fen, and woody vegetation over time (Tansley, 1939). This process is slowed by extensive livestock grazing, and in the prehistoric past by grazing by large wild herbivores. However, if sediment inputs and nutrients are both low, the pond may persist as a bog or fen (Walker 1970). Alternatively, pond sediments may accumulate to the point where the water is so shallow it dries up for long periods. When sediments are exposed to the air, they oxidise and decompose much more quickly than in water. This may effectively halt the infilling process, allowing some temporary ponds to persist for thousands of years (Wood et al. 2003, Williams et al. 2001).

Data from surveys suggests a continuum of pond types, ranging from permanent to seasonal ponds, with approximately 70 % being permanent, 25 % semi-permanent, and 5 % seasonal in the British lowlands (Biggs & Williams, 2023 p191). Temporary ponds, common in drier regions, have garnered significant research attention, with evidence suggesting they can persist for thousands of years (Wood et al., 2003; Williams et al., 2001). Paleoecological research suggests that lowland, floodplain ponds in a non-impacted human landscape persist between 2,000 and 3,000 years (Biggs & Williams 2024, p298).

Numbers and area of ponds

Estimates of pond number and area worldwide vary, partly due to a lack of agreement over the size boundaries of a 'pond'. Current estimates are 100 - 300 million lentic waterbodies worldwide between 0.1 and 10 ha, accounting for around 30 % of global standing water area, or between 1 and 2 million km ^{2}. (Downing et al. 2006, Verpoorter et al. 2014, Holgerson & Raymond 2016). These estimates all exclude large numbers of water bodies less than 0.1 ha, which may number 3.2 billion (Downing et al. 2006). The UK Countryside

Survey (2007) found that 70 % of water bodies were between 25 m^2 and 400 m^2 (Carey 2008). Whatever the exact number, these small water bodies make up a significant proportion, in number and extent, of total surface freshwater.

At the most recent count, there were an estimated 478,000 ponds in Great Britain between 25 m^2 and 2 ha (Carey 2008). This was a 38 % increase of ponds on the historic low recorded in 1990, although it is still a fraction of the 1.2 million ponds in 1880 (see Figure 1.2). Additionally, Carey (2008) recorded a high rate of pond turnover (18,000 ponds lost and 70,600 gained), so the current number is likely different. In a typical lowland British landscape (the Severn Vale), a recent study calculated that 58 % of ponds were lost between 1900 and 2019 (Clarke et al. 2022). It should be noted that these pond numbers do not include garden ponds, which are predicted to number between 2.5 and 3.5 million (Davies et al. 2009).

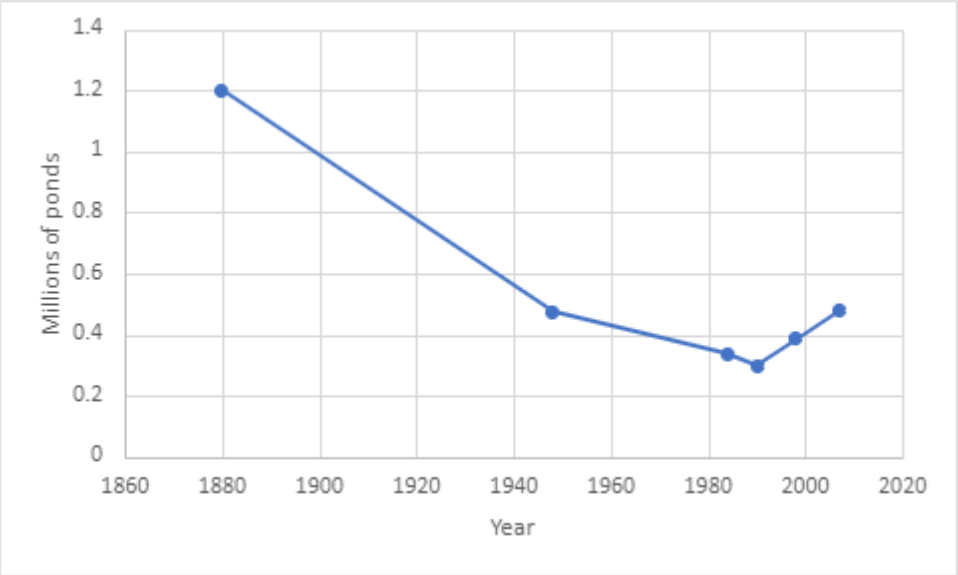


Figure 1.2: Numbers of ponds in Great Britain (England, Wales and Scotland) over the past 120 years. Adapted from Biggs et al. (2005) with data from the Countryside Survey 2007 (Williams et al. 2010).

Pond communities

The margins of ponds share features with larger freshwater wetlands. These include saturated soils which create anaerobic conditions which the biota must adapt to, and the presence of macrophytes (Van der Valk 2012). These margins, which often vary in their dryness and extent at different times of the year, are often the richest part of the pond

community and interact extensively with the more purely aquatic environment (Biggs & Williams 2024 p249).

Macroinvertebrates in ponds include water snails and bivalves (Mollusca*)*, shrimps and water slaters (Crustacea*),* water beetles (Coleoptera*)* water bugs (Hemiptera*),* flatworms (Platyhelminths*)*, mayflies (Ephemeroptera*),* caddisflies (Trichoptera*)* and dragonflies and damselflies (Odonata*).* The main differences in the invertebrate assemblages between ponds and other freshwaters are that ponds have a higher proportion of water beetles – between a third and a half of the macroinvertebrate species typically – as well as a greater proportion of water bugs and dragonflies and damselflies (Biggs & Williams 2024*).*

431 species in these surveyed groups have been found in ponds (Biggs et al. 2005). Other invertebrates include true flies (Diptera) and segmented worms (Oligochaetes), which have not been surveyed to the species level in Britain but are expected to be diverse groups (Biggs & Williams 2024 p347). Microinvertebrates (invertebrates which are not visible to the naked eye and cannot be sampled with a net) have also not been surveyed to species level, but groups include water fleas (Cladocera), Copepoda and seed shrimps (Ostracoda), Rotifera, mites *(*Arachnida*),* Nematode worms and others such as *Hydra* and Gastrotricha.

Many invertebrate species exist as adults in the aquatic environment (e.g., water snails), whereas many others live aquatically only at the larval (e.g., caddis flies) or nymph stage (e.g., Damselflies). Some breathe using gills, whereas others use air, for which there are many behavioural and morphological adaptations (e.g., breathing tubes, diving bubbles) to access. Different species or life stages of the same organism may inhabit the benthos, water column, submerged and emergent macrophytes, sediment or the water surface (Burton 1977).

Around 250 species of higher plants and ferns can be found in ponds in the UK. Large plants that live in water or flooded soils (macrophytes) may be divided into three broad groups: emergent plants (roots in flooded soil, leaves in air), submerged aquatic plants (roots in sediment, leaves under water) and floating leaved aquatic plants (roots in sediment or water, leaves on water surface). For all freshwater habitats in the UK, emergent plant species make up the greatest proportion of the plant community, at around 80 %. Floating-leaved plants are more common in turbid ponds, whereas submerged plants are an indicator of clean water ponds (Biggs & Williams 2024).

There are seven native amphibian species in Britain, all of which breed in ponds. Some are common and widespread (e.g., *Rana temporaria, Lissotriton vulgaris)* whereas others are

vanishingly rare (e.g., *Epidalea calamita*). The Great Crested Newt (*Triturus cristatus)* is a strictly protected species and is a key focus of conservation schemes and research (Lewis et al. 2016). Around 15 species of freshwater fish inhabit British ponds and were present in 60 % of ponds in the first national pond survey (Biggs & Williams 2024) Ponds are an important habitat for many bird and mammal species and the grass snake *(Natrix helvetica).* For instance, a recent study found that open canopy farmland ponds supported 18 times the abundance of emergent insects than overgrown ponds, which in turn was linked to a greater abundance and richness of bird species (Lewis-Philips et al. 2020).

Algae are a large and diverse group of organisms in ponds, and highly important for ecosystem functioning. The term “algae” covers several branches of the phylogenetic tree of life and refers to any simple organism with a photosynthetic pigment. This includes the eukaryotic algae - green algae (Chlorophyta*)*, diatoms (Bacillariophyta*),* golden-brown algae (Chrysophyceae*),* Dinoflagellates*,* Euglinida and others, but sometimes prokaryotic algae (Cyanobacteria) as well. They may be found suspended in the water column (phytoplankton), growing on surfaces such as macrophytes, rocks or sediment (periphyton), or in large mats (metaphyton). Large green algae such as the stoneworts (Charophyceae*)* and filamentous green algae such as *Spirogyra* and *Oedogonium* are also common in ponds.

In wetland ecosystems, it is estimated algae make up around 20 % of net primary productivity (Van der Valk 2012), and they’re also an important food source for grazing invertebrates and detritivores. Due to their large diversity and small size, few studies attempt to identify this group to species level. However, research from the first half of the 20th Century suggests around 100 species of algae can be found in a pond in Great Britain (Biggs & Williams 2024, p299), and more recent research found a similar number (although only counting phytoplankton, Smith et al. 2005).

Other micro-organisms are so little studied in ponds that they are barely mentioned in Biggs & Williams 2024, the most recent monograph on UK ponds: in over 600 pages, “fungi” has 12 mentions, “bacteria” 21 and “protozoa” 13. In wetland habitats in general, it is thought that most of the microbiota are shared with other aquatic habitats, whereas the macrobiota are more specialised (Van der Valk 2012 p45). Bacteria may be found free in the water column, or on surfaces and in the soil/sediment. Most are thought to be non-photosynthetic and probably saprophytic. Anaerobic bacteria and those which respire using methanogenesis are common in the anoxic environments of wetlands, such as the sediment, and are crucial for the cycling of carbon, nitrogen, and sulphur (Van der Valk 2012, Chapter 6). Fungi

associated with wetlands may be obligate aquatic, have an aquatic life stage, or may be terrestrial but tolerate flooding. Most are found attached to a surface, the only exceptions being yeasts and yeast-like fungi in the water column. Fungi are important decomposers but most often are aerobes, so are absent from anoxic areas of the wetland environment, which accounts for the low rates of decomposition in ponds (Van der Valk pp49-50).

Biodiversity importance of ponds

Temperate ponds support an impressive array of freshwater species. In fact, ponds may be the most biodiverse type of freshwater habitat in Great Britain and Europe (Williams et al. 2004, Davies et al. 2008). An overall comparison between pond and river macroinvertebrate databases in Britain found that ponds supported 10 % more species than rivers despite representing three times fewer ponds than river sites (Biggs et al. 2005). Ponds in the UK also support significantly more rare species than streams, rivers, or ditches, as evidenced by the comparison of species rarity indices for the four habitats nationwide (Biggs et al. 2007).

Williams et al. (2004) compared plant and macroinvertebrate diversity in rivers, streams, ponds, and ditches over an 80 km^2 lowland agricultural landscape in Oxfordshire. Ponds had lower alpha species richness than rivers, but over the entire landscape, ponds supported 71 % of all species and rivers 60 %, despite ponds covering the smallest area of the five different types of waterbodies. A similar pattern was found in a study incorporating other agricultural landscapes in Shropshire, France, Denmark, and Germany (Davies et al. 2008); ponds had the highest gamma diversity for macrophytes in all five locations.

Several smaller temperate waterbodies often have greater species richness than the equivalent area of a single larger waterbody. This configuration has been found in studies worldwide, for example, in Swiss ponds for six different groups of aquatic fauna and flora (Oertli et al. 2002), algae in Hungarian ponds (Bolgovics et al. 2019), littoral invertebrates in mountainous ponds in Spain (Martinez-Sans et al. 2012), and rare wetland plants in New Zealand (Richardson et al. 2014). There has been limited research on other habitats (e.g., the tropics).

Therefore, ponds refute the ecological principle that a single large habitat patch holds more species than several small patches (also known as the “SLOSS” principal, Diamond 1975, Pfeiffer 2017). However, mounting evidence also contradicts this hypothesis in other habitats. Recent reviews have found that most studies report positive biodiversity responses in habitats with more fragmentation per se (76 % of studies, Fahrig 2017) and greater

species richness in several small patches than in a single large patch of the same area (52 %, Fahrig 2020).

Why do ponds exhibit this pattern? Biggs et al. (2005) and Davies et al. (2008) put forward a hypothesis based upon habitat heterogeneity: ponds have a smaller catchment size than rivers and lakes, and are strongly affected by local conditions, and so there is more variation between ponds' physical and chemical parameters compared to larger waterbodies. Lakes and rivers have a larger catchment, and because of the mixing of water, the conditions across the catchment are averaged across the entire lake or river. The result is many ponds of varied condition compared to fewer large water bodies of similar conditions (Williams et al. 2004, Angelibert et al. 2004). However, Scheffer et al. (2006) proposed a hypothesis based on species interactions: the absence or low abundance of fish in smaller European lakes has a positive effect on macrophyte, and in turn invertebrate and bird, abundance.

Despite differences in species richness, different freshwater habitats in the UK showed a significant overlap in their community composition. A comparison of nationwide invertebrate databases from rivers (RIVPACS, Wright et al. 1996) and ponds (National Pond Survey 1996) found that 30 % of the species were found only in ponds, 23 % in rivers, and 43 % in both habitats (Biggs & Williams 2024 p101). In plants, the overlap is similar (Williams et al. 2004; Biggs & Williams 2024 p99). The overlap is even greater with lakes, with 70 % of species found in both habitats (Wright et al. 1996)

These findings highlight the importance of conserving all freshwater bodies in an area as an integrated network rather than focusing on individual freshwater habitat types in isolation. After all, freshwater species evolved in a wetland environment which was a "blur of interconnected lakes, rivers, ponds and swamps" (Sayer 2014). Similarly, the areas of land in between those freshwater areas, or the 'matrix,' are also highly important to conserve freshwater species, as I shall explore in more detail in section 3.

Pond ecosystem services

In addition to their high biodiversity value, ponds provide other important ecosystem services, such as small-scale water storage, pollution control and carbon cycling (Biggs et al. 2017). There is extensive evidence from around the globe to show that natural and artificial ponds can retain nutrients and reduce pollution loads at a catchment level (Tournebize et al. 2016). Ponds without surface water connections to other waterbodies can still impact downstream flow, buffering stream variation (Golden et al. 2015). Ponds are vital water

stores in deserts, plains, and areas of high agricultural impact (Bichsel et al. 2016), and a key biodiversity resource in urbanised landscapes (Hill et al. 2017).

The impact of small waterbodies on the global carbon cycle is an area of active research. Downing (2008, 2010) suggested that given their ubiquity, their high levels of organic sediments and high biogeochemical activity, small lakes and ponds could bury more carbon than the oceans. The last decade has seen an increase in measuring carbon fluxes of ponds, with multiple evidence that they are a significant source of methane and carbon dioxide (Holgerson & Raymond 2016), particularly in boreal and arctic regions (Kuhn et al. 2018). Aquatic ecosystems may account for half of all global methane emissions, with 37 % of “lake” emissions coming from small ponds <0.1 ha (Rosentrater et al. 2021). On the other hand, Taylor et al. (2019) calculated that small ponds have high carbon burial rates, over 20 times that of woodlands or grasslands.

Threats to ponds

Ponds have experienced a significant decline in both numbers and quality. A study in the UK in 2007 revealed that 80 % of ponds were rated as poor or very poor in quality, with a 20 % reduction in plant species richness compared to 1996 levels (Williams et al., 2010). Even minimally impaired ponds, primarily located in nature reserves and managed for species conservation, demonstrated a decline in macrophyte species richness (Williams, 2018).

A 2018 document listed a number of threats to ponds in the UK (see Table 1.1), which could be applied to ponds in other areas (Cereghino et al. 2007, Biggs et al. 2017). I will examine each of these in more detail below:

Table 1.1: Broad threats to freshwater environments, from Dudgeon et al. (2006) and examples of their impact in ponds, from CaBA (2018)

Threat category (from Dudgeon et al. 2006)	*Examples relevant to ponds (from CaBA 2018)*
Water pollution	Nutrient enrichment; air pollution; other pollutants such as pesticides or heavy metals
Flow modification	Drainage and infilling, groundwater abstraction or land drainage

Habitat degradation	Terrrestrialisation, changes in land use practices
Species invasion	Invasive and non-native species
Over-exploitation	Use for recreation
Climate change	Climate change

Of these threats, pollution, especially nutrient pollution (nitrogen and phosphorus), poses a significant risk to ponds, echoing concerns for freshwater habitats globally (Reid et al., 2019; Dudgeon et al., 2006; IPBES, 2019). High nutrient levels, fertiliser application and land use intensity were all associated with decreased macrophyte species richness in British lowland ponds (Biggs & Williams, 2023 p235). Eutrophication, linked to elevated nutrient levels, negatively impacts macrophyte and invertebrate richness in shallow lakes and ponds worldwide (Sondegaard et al., 2010; Philips et al., 2016; Stefanidis et al., 2019).

Terrestrialisation, characterized by sediment infilling and woody plant encroachment, is another critical threat to ponds. It is accelerated by human activities leading to more sedimentation, and the disappearance of wild and domesticated large herbivores which keep ponds open (Robinson & Sutherland, 2002). The filling rate of ponds due to this process ranges from 1-3 cm per year (Williams et al., 1998), potentially leading to complete infilling within 40-120 years (Biggs & Williams, 2023). Shade negatively correlates with species richness, but its impact is confounded by land use intensity, as most shaded ponds in the countryside are within arable fields (Williams et al., 1998, 2010, 2018). Intensive grazing around ponds may lead to increased sediment load in the pond, and decreased vegetation complexity (Declerck et al. 2006).

Invasive species, though a major threat to freshwater habitats, have a limited impact on ponds in the UK, with only 10 % of ponds hosting non-native species (Williams et al., 1998, 2010). Over-use by people and dogs, and over-stocking with fish and ducks may have strong negative effects on the plant community, eliminating submerged macrophytes, although at a national level, these are thought not to have a significant effect (Williams et al. 2010).

Climate change is anticipated to affect ponds, altering hydrological regimes and species distributions due to changing temperatures and precipitation patterns (Reid et al., 2019). Ponds, due to their small size and catchment, may experience hydrological and thermal shifts more rapidly than larger water bodies (Matthews, 2010; Clarke, 2009), although effects may be buffered by pond species' wide thermal tolerance (Riley et al. 2018).

One threat not listed in either paper is the lack of freshwater habitat connectivity. In a typical lowland landscape, pond density has declined from 7.5 per km ^{2} to 4.3 per km ^{2}, and the average distance between ponds has increased by 25 m (Clarke et al. 2022). This may not sound like much, but it could be a significant threat to ponds, as the decrease in the number and density of water bodies in the landscape greatly reduces the chance of successful colonisation events. Alarmingly, the full effects of the loss of connectivity may not be seen yet due to extinction debt (Semlitsch et al. 2017). In a recent study of a Leicestershire farming area, 10 % of plant species were only present in one pond (Williams et al. 2020).

There are indications that connectivity is important for pond biodiversity, although its relative importance compared to other variables is unclear. For instance, Oertli et al. (2002) found that connectivity (number of water bodies in a 1 km radius of the pond) accounted for 9 % of the variation in amphibian richness between ponds (as a comparison, shade accounted for 9 % of the variation in Odonata richness, and eutrophication level 16 % of the variation in aquatic plant richness). The 2007 Countryside Survey found that close proximity to other freshwater (<100 m) was correlated with higher plant species richness (Williams 2010). Macrophyte richness in ponds in Northern Italy decreased significantly with increasing isolation, and, after pond area, was the most important explanatory variable for richness (Bolpagni et al. 2020). See section 3 for a more thorough examination of this topic.

Pond conservation

Traditionally, pond conservation has taken one of two approaches: pond creation or pond restoration. Pond creation aims to increase pond numbers and density in the landscape and to create ponds with high water quality by siting ponds in areas of semi-natural habitat or buffering anthropogenic impacts (Pond Conservation 2009). In addition, pond creation increases the number of early successional ponds but does not remove late-succession ponds (Biggs & Williams 2024, Chapter 12). As there are pond organisms adapted for different stages of succession, this should increase overall biodiversity (Hassall et al. 2012).

Manmade ponds can be just as biodiverse as naturally created ponds. For instance, ponds in Pinkhill Meadow (a complex of 40 ponds, scrapes and pools created in 1990), have on average 40 wetland plant species and 47.5 invertebrate species (Biggs & Williams 2024, p632), much higher than the average richness in either wider countryside ponds (Williams et al. 2010) and minimally impaired ‘reference’ ponds (Williams 2018). Another recent study of Irish ponds found that created ponds in coastal agricultural grasslands had higher macroinvertebrate richness than natural ponds in nearby sand dunes (Reyne et al. 2021), despite artificial ponds being, on average, smaller.

Pond creation increases macrophyte diversity at the catchment level. Freshwater habitats (streams, ditches, and ponds) of three small lowland agricultural catchments were surveyed over nine years for macrophyte richness and rarity. The control catchment displayed a 1 % decline in macrophyte richness per year. Nature-based interventions (e.g., bunded streams) reversed this trend, but the creation of ponds increased macrophyte catchment richness by 26 % and more than doubled the number of rare plant species (Williams et al. 2020).

Pond restoration efforts include removing trees and shrubs from the pond margin and dredging sediment from the pond to revert the pond to an earlier successional stage (Sayer et al. 2012). There is evidence of this approach providing benefits to biodiversity: in a study of 40 managed and unmanaged farmland ponds in Norfolk, managed ponds had significantly higher macrophyte and invertebrate diversity, and this tended to peak 3-5 years after management (Sayer et al. 2012). Recent research on the same ponds found an 18-fold higher abundance of insects in managed ponds, which in turn supported a higher abundance of farmland birds (Lewis-Phillips et al. 2019, 2020). However, shade-adapted species should be considered when planning restorations (Natural England, 2016).

Considering the conservation of ponds at the habitat-scale only may not be sufficient. For instance, Sawatasky et al. (2019) found that amphibian occupancy of ponds in agricultural landscapes was far more strongly controlled by natural habitat in the surrounding 1 km radius than the presence or absence of “buffer” strips around ponds. Landscape-scale approaches are required, such as the “catchment-based approach”, which considers all waterbodies in a catchment, and different users, and uses, of the land and water. There are currently over 100 catchment partnerships in the UK, which have brought social and economic benefits, but clear benefits to biodiversity are yet to be seen (CaBA 2018). Pond creation is promoted as a “nature-based solution” to both climate change and biodiversity loss (Cuenca-Cambronero et al. 2023), although the contribution of manmade ponds to carbon emissions is still unclear (see above).

Section 2: eDNA metabarcoding in freshwater ponds

An introduction to environmental DNA and metabarcoding

a. A short history

In the past five years, the use of the terms ‘environmental DNA’, ‘eDNA’ and ‘metabarcoding’ have exploded in scientific literature, and its use in the general media is rising (see Figure

1.3). eDNA has been hailed as a tool which is “transforming how we survey plant and animal communities” (Deiner et al. 2017) and as a “key component of the ecologists’ and environmental managers’ toolbox” (Taberlet et al. 2018).

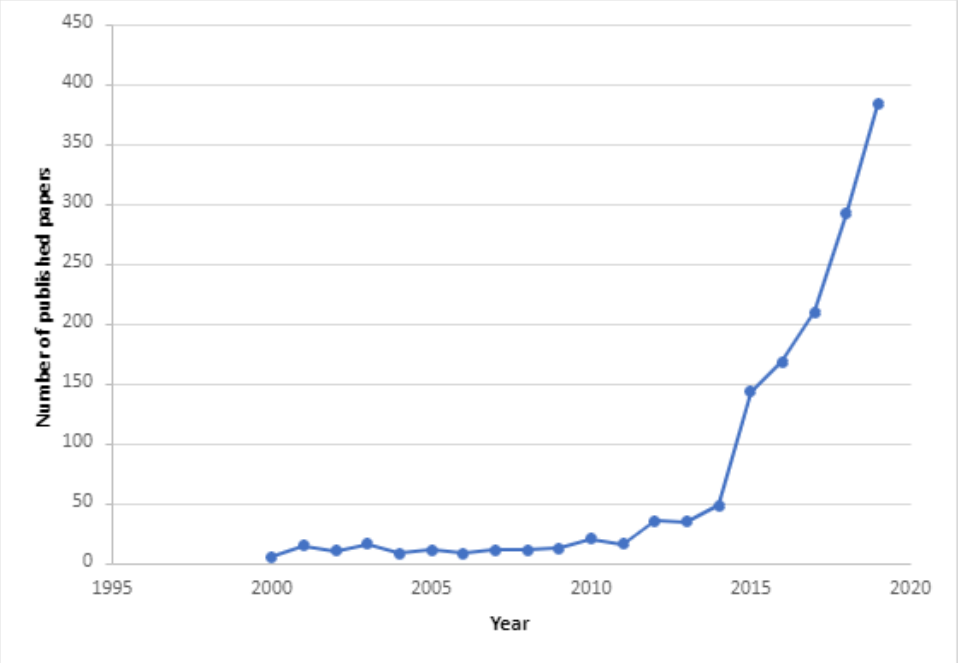


Figure 1.3: Number of papers published each year focussing on eDNA. Data returned from a search of Web of Science, Title = “eDNA” OR “environmental DNA” OR “metabarcoding”

Environmental DNA, or eDNA is “a complex mixture of genomic DNA from many different organisms found in an environmental sample” (Taberlet et al. 2018). Environmental samples include water, soil, sediment, or even air (Taberlet et al. 2018, Harper et al. 2019a, Deiner et al. 2017). The DNA may be intracellular or extracellular in origin, and may have originated from skin cells, mucus, saliva, scales, gametes, pollen, deceased remains or many other sources (Harper et al. 2019, Ruppert et al. 2019). There is sometimes a distinction drawn between intracellular and extracellular eDNA, with some authors excluding the former from the definition of eDNA (Barnes & Turner 2016). However, this would rule out all microbial DNA in environmental samples, so I do not make this distinction in my research.

Scientists have been aware of the ‘soup’ of DNA in the environment for many years, and eDNA from micro-organisms was sequenced as early as 1987, using cloning and Sanger sequencing (Ogram et al. 1987). The commercialisation of high throughput sequencing (HTS, also called next generation sequencing, NGS) in 2005 enabled researchers to sequence millions of DNA fragments simultaneously, and since then, there has been a rapid increase in studies using environmental DNA. Initially, microbial communities were

sequenced (Giovannoni et al. 1990), followed by ancient communities from sediment (Willerslev et al. 2003), and finally eDNA from macro-organisms (e.g., Ficetola et al. 2008).

In the last decade, there has been a significant increase in the frequency and scope of eDNA studies. Techniques based on the analysis of eDNA have been used to detect single species, entire communities, and even genomes and functional genes of those communities. eDNA has been applied to early detections of invasive and rare species, biodiversity monitoring and diet analysis, amongst other things (for reviews, see Deiner et al. 2017, Ruppert et al. 2019, Cordier et al. 2020, Beng & Corlett 2020). Its ability to sample habitats in a non-invasive and non-destructive manner, and its speed, efficiency and scalability give it the potential to revolutionise how we understand the natural world (Baird & Hajibabaei 2012).

b. An introduction to the method

No matter what the application, eDNA methods broadly follow the following steps: the eDNA is captured, extracted, amplified, sequenced, and then identified or assigned (Harper et al. 2019). The method will differ in detail depending on the environmental sample, the target and the research aim. For example, when sampling eDNA from freshwater, a water sample may be collected and subsequently filtered through a very fine membrane (<2 µm). The captured eDNA is then extracted from the filter papers, usually using a combination of physical and chemical extraction methods. On the other hand, eDNA does not need to be filtered from soil or sediment and can be extracted from these samples directly. The captured eDNA is then amplified, most often by using the polymerase chain reaction (PCR). The key tool in this process is the PCR primers: short, single-stranded pieces of DNA which bind to the denatured DNA in the sample, and to which, subsequently, nucleotides are added to elongate and so amplify (make many copies of) the DNA. Primers are targeted so they amplify a single region of DNA of interest. These copies of the DNA sequence of interest are called “amplicons” (for a more detailed description of the method, see Taberlet et al. 2018.) Two main approaches are the targeted approach using single-species primers, or a non-targeted approach using universal primers.

Single-species primers bind to and amplify a gene sequence which is highly conserved for every member of that species but not shared by any other species in that environment. If the target gene is present in the sample, PCR amplifies the gene and then quantitative approaches give a signal that it is present. From these results, the presence or absence, and even the abundance, of species in an environment can be inferred (Thomsen et al. 2012, Spear et al. 2020)

PCR approaches are simple and sensitive, but they require knowing which species are in your environment *a priori.* On the other hand, if you wish to be able to identify multiple unknown species or taxa within a community, universal primers are often used. These primers are (ideally) designed so they bind to two conserved regions of the genome, shared by all target taxa, which flank a variable region which is different for every individual taxon of interest (see fig 2, below). The variable region is called a 'metabarcode', and the process of identifying multiple taxa within a sample 'metabarcoding'. It is carried out using a combination of PCR to amplify sequences, and then high-throughput sequencing (HTS) of the amplicons. HTS can sequence millions of different DNA amplicons within a single sequencing 'run', allowing the identification of multiple taxa within a sample, and between samples, simultaneously.

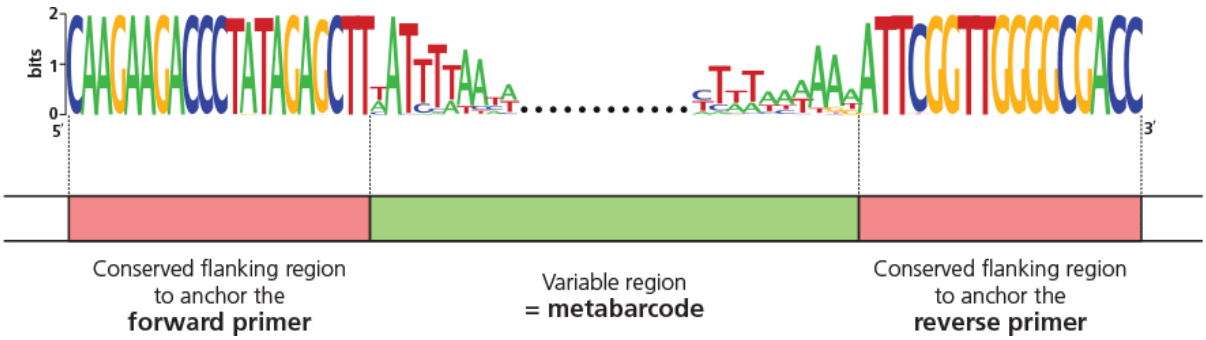


Figure 1.4: An example of a primer pair and variable metabarcode region, in this case targeting suborder Lumbricina (earthworms). The height of the nucleotide indicates how conserved it is at that position. From Taberlet et al. 2018.

The primers used for metabarcoding are always less specific than those used for species identification, but they may vary in their specificity; for instance, they may be designed to amplify entire domains of life (e.g., Bacteria or Eukaryota) or lower phylogenetic levels (e.g., copepods, diatoms, or fish) (see Taberlet et al. 2018, Appendix 1 for examples). Typical genes targeted by metabarcoding primers include 16S rRNA for bacteria, 18S rRNA for eukaryotes and mitochondrial Cytochrome Oxidase I for animals. It is a challenge (and perhaps impossible) to design a primer pair which equally amplifies all the species in the target taxon (breadth) to a level of detail which can distinguish between species (resolution). Often, a given primer pair provides a compromise between taxonomic breadth and resolution (Deagle et al. 2014).

The raw data from HTS requires extensive bioinformatic processing to reduce noise and turn the output into something which can be used for taxonomic classification. This can involve

many steps, including ‘cleaning’ data to remove sequencing errors, aligning forward and reverse reads, clustering and de-replicating data. Many tools or ‘pipelines’ can be used for these steps, such as QIIME 2 (Bolyen et al. 2019) or DADA2 (Callahan et al. 2016). Finally, the clustered sequences are compared to reference databases to assign taxonomy. There are many reference databases (Taberlet et al. 2018, Chapter 3), which vary in their geographic and taxonomic coverage, completeness, curation and how often they are updated.

c. Current challenges

There are many sources of bias that determine whether eDNA analysis will detect a specific taxon in a certain environment. Firstly, there is the ecology of the eDNA itself: differences in its origin, state, transport and degradation rate will influence how much of the eDNA is collected and captured (Barnes & Turner 2016). Secondly, there are the field methods deployed, such as the sampling regime for collecting the DNA, and any anti-contamination protocols (Dickie et al. 2018). Thirdly, laboratory methods will have an effect: some eDNA extraction methods may perform better than others, and some taxa are preferentially amplified by primers over others (“primer bias”) (Bradley et al. 2016, Lear et al. 2018). Finally, there are *in silico* methods, such as the post-sequencing processing of the data (“bioinformatic pipeline”) and the reference databases used (Tapolczai et al. 2019).

There have been calls to move away from purely technical eDNA studies towards clear hypothesis testing (Prosser 2020), and it’s generally agreed that more standardisation amongst eDNA studies is needed (Lear et al. 2018). Many reference databases are incomplete and biased in their taxonomic and geographic coverage (Hestetun et al. 2020, Willerlev & Thomson 2015), which can significantly influence the results of biodiversity assessments using eDNA (Schenekar et al. 2020). Some have gone as far as to say some environmental DNA studies are unscientific, suffering from poor experimental design with low reproducibility and repeats (Zinger et al. 2019). To date, there is a poor representation of eDNA studies in the tropics compared to temperate regions, but terrestrial, freshwater and marine environments are all relatively equally represented (McGee et al. 2019).

d. Future applications

There are many and various potential future uses of eDNA. Emerging areas of research include air and water quality monitoring (Ahmed et al. 2019, Banchi et al. 2018), disease detection (Miaud et al. 2019) and plant-pollinator interactions (Oliver et al. 2021). A fundamental issue is how eDNA metabarcoding can be used as an ecosystem and

biodiversity monitoring tool, either replacing traditional taxonomic monitoring or by providing additional insights and metrics (Baird & Hajibabaei 2012, Deiner et al. 2017, Pawlowski et al. 2018, Ruppert et al. 2019, Cordier et al. 2020). This is explored in more detail in section 3.

Another avenue of study is employing multiple different primer pairs on eDNA collected from the same sample to monitor entire communities, for instance, bacteria, eukaryotic microbes, metazoans and vertebrates simultaneously (e.g., Zhang et al. 2020). An active research question is whether eDNA "read abundance" in metabarcoding can be used to infer the abundance of the taxon or taxa in the environment (Machler et al. 2020), with current thinking that there is a weak positive relationship between biomass and read abundance (Lamb et al. 2018)

Moving away from simple detection and "who lives where" studies (Prosser 2020), eDNA metabarcoding has the potential to test almost any ecological hypothesis that requires sampling of communities. For instance, it may be used to construct food webs and other ecological networks, which can subsequently be used to investigate the effects of disturbance on ecosystem resilience (e.g., Evans et al. 2016), or for determining assembly rules of communities (e.g., Stoof-Leichensring et al. 2020).

In the following sections I shall review the use of this technology in freshwater ecosystems and specifically ponds.

eDNA studies in freshwater environments

Freshwater has been a very productive area for eDNA research, although challenges remain.

a. Singe species detection: rare, invasive & commercial species.

Detection of rare and invasive vertebrates using eDNA in freshwater was one of the first applications of the method. Ficetola et al. (2008) used the method to detect the invasive American bullfrog in aquaria and natural wetlands in France, proving that false positives and negatives could be discriminated. More recently, it's been used to detect rare amphibian species in tropical environments and swamps (e.g., Eiler et al. 2018 and Goldberg et al. 2018), with suggestions that in some cases it may be more effective than traditional monitoring. An eDNA assay was developed to detect Great Crested Newts, a protected species in the UK for which monitoring is legislated (Biggs et al. 2015) and is now used for nationwide monitoring by governmental bodies (Natural England 2019).

Single species detection was quickly expanded from amphibians to invasive fish species such as the Asian carp (Jerde et al. 2011). Now, there are ubiquitous studies on numerous continents using eDNA to detect rare or invasive fish species: recent examples include the detection of a non-native salmon in Patagonia (Chalde et al. 2019), endangered native fish in Japan (Jo et al. 2020), and to assess the success of fish reintroduction in Germany (Riaz et al. 2020).

Reptile and mammal species have been successfully detected, such as the invasive Burmese python in Florida (Piaggio et al. 2014) and North American river otters (Padgett-Stewart et al. 2016). The method has been expanded to some invertebrate species, such as the signal crayfish (Robinson et al. 2018), invasive mosquito species across water bodies in Europe (Schneider et al. 2016), and zebra mussels in the Great Lakes (Gingera et al. 2016).

Plant detection came a little later, with assays developed in 2015 (Scriver et al. 2015) and used to detect invasive *Egeria densa* in Japan (Fujiwara et al. 2016). Current research is in a similar vein, for instance, detecting invasive Canadian pondweed in Norwegian lakes and streams (Angles d'Auriac et al. 2019).

When detecting single species using qPCR, multiple studies have now found that the concentration of eDNA is linked to the biomass and therefore the abundance of the species of interest e.g., plants (Matsuhashi et al. 2016), and fish (De Muri et al. 2020) Due to this relationship, eDNA has been suggested as a tool to monitor populations and abundance, for instance of commercial fish species (e.g., Lacoursiere-Roussel et al. 2016a).

b. Whole communities

i. Vertebrates, macroinvertebrates and plants

Thomsen et al. (2012) carried out the first whole-community 'metabarcoding' study on macro-organisms, targeting two amphibian species, otter, spined loach, a dragonfly species and a crustacean in ponds. However, this study used specially adapted primers for each species and qPCR, so properly is several single-species assays combined rather than 'true' metabarcoding.

Characterising species composition, richness and abundance of macroinvertebrates in freshwater using eDNA metabarcoding is an attractive prospect because this group is frequently used as a measure of water quality, for instance under the European Union Water Framework Directive (Davy-Bowker et al. 2006). Studies have focussed on developing primers which reliably amplify across macroinvertebrate taxa (Hajibabaei et al. 2011, Elbrecht & Leese 2017), assessing whether the relative abundance of taxa can be assessed

(Elbrecht & Leese 2015), and directly comparing the morphological and metabarcoding approaches (many, e.g., Elbrecht et al. 2017, Deiner et al. 2016, Fernandez et al. 2018, Serrana et al. 2019). The relevance of these studies to biomonitoring is covered in section 3.

Vertebrate metabarcoding is a rapidly expanding field of study. It appears that fish communities can be detected equally well by metabarcoding as traditional methods such as electrofishing or netting, both in lentic (Hanfling et al. 2016, Evans et al. 2017, Civade et al. 2016) and lotic (Olds et al. 2016, Shaw et al. 2016, Nakagawa et al. 2018) ecosystems. Expanding from single amphibian species, several studies have used the metabarcoding approach to detect multiple amphibian species (e.g., Lopes et al. 2017), or even both fish and amphibian species simultaneously (Valentini et al. 2016).

Excitingly, eDNA from water has shown the potential to detect mammal species, not only aquatic species but also terrestrial species which interact with that water source. For instance, Ishige et al. (2017) detected six endangered species from salt-lick water in Bornean rainforest, and mammal species have also been detected from Namibian watering holes (Seeber et al. 2019) and upland streams in the UK (Sales et al. 2019).

Metabarcoding studies have seen a comparative lag in focus on plants, partly due to lack of agreement over a universal barcode for plants: a selection of genes targeted by primers include ITS2, matk, rbcL and the trnL P6 (Fahner et al. 2016, Dormontt et al. 2018). Studies tend to focus on eDNA in sediment rather than in the water column, and its predominant use has been to reconstruct ancient plant communities from lake sediment, often alongside pollen analysis, and often in high latitude regions (e.g., Jorgensen et al. 2012, Parducci et al. 2018, Willerslerv et al. 2014), but there are some studies in tropical regions (e.g., Boessenkool et al. 2014).

Studies attempting to sample the contemporary plant community are rare, but this has been attempted in lakes, from surface sediments (Alsos et al. 2018), and water (Drummond et al. 2021), and in floodplain wetlands (Shackleton et al. 2019). Work is ongoing to develop specific aquatic plant primer sets (Coghlan et al. 2020), including for specific taxa such as *Potamogeton* (pondweeds, Kuzmina et al. 2018)

ii. Microbes, microfauna, algae & fungi

Metabarcoding provides a valuable tool for studying smaller organisms which are more costly and time-consuming to survey in freshwater ecosystems, including microorganisms (prokaryotic and eukaryotic), algae, fungi, and micro- and meiofauna. eDNA metabarcoding has been used extensively to study these organisms in lakes, rivers, and wetlands, ponds

have been relatively overlooked due to a general scientific neglect of smaller habitat patches and exclusion from legal frameworks (Riva & Fahrig 2022, Biggs et al. 2017).

Before high-throughput sequencing, bacterial community characterization began with molecular techniques. The first metabarcoding studies targeted bacteria in aquatic environments (Giovanni et al., 1990; Ward et al., 1990). HTS revolutionized sequencing rates, but the immense diversity of prokaryotes still challenges practical diversity calculation (Shafi et al., 2017). The Earth Microbiome Project (EMP) initiated in 2010 aimed to sequence microbial communities across all major biomes, including freshwaters (Gilbert et al., 2010).

In river water samples, the bacterial phyla Proteobacteria, Bacteroidetes, Actinobacteria and, to a lesser extent, Verrucomicrobia are the most abundant (Cruaud et al. 2020, Li et al. 2020, Gweon et al. 2019, Doherty et al. 2017, Read et al. 2015). In river sediment samples, Proteobacteria and Bacteroidetes were still abundant, but there are greater proportions of Acidobacteria, Chloroflexi and Firmicutes (Yuan et al. 2023, Liu et al. 2022, Wu et al. 2019, Liu et al. 2018). (Lv et al. 2014). A similar pattern in water and sediment is found in lakes (Jiao et al. 2021, Liu et al. 2020, Nakatsu et al. 2019, Ruuskanen et al. 2018, Zhang et al. 2019) and in wetland sediments (Lv et al. 2014).

Although prokaryotic compositions are similar at the phyla level, compositions can still alter with finer-scale changes in the freshwater environment e.g., with increasing dendritic distance in a river system (Read et al. 2015), and within 30cm of lake sediment (Wurzbacher et al. 2017a). On an even smaller scale, bacteria have exhibited community turnover within 30 m in a small freshwater pond (Lear et al. 2014). Bacteria functions are thought to be phylogenetically conserved (Isobe et al. 2019), making taxa good indicators of different ecosystem functions (Urakawa & Bernhard 2017), although freshwater bacteria lineages within a single genus can also have very different ecological preferences (Nuy et al. 2020).

Compared to bacteria and archaea, studies of small eukaryotes using the 18S rRNA gene region yield less consistent results. Lake water samples reveal a wide range of dominant taxa: Ochrophyta, Dinoflagellata, Chlorophyta, Ciliophora and Cryptophyta were commonly reported, alongside Arthropoda (Debroas et al. 2017, Mikhailov et al. 2018, Banjeri et al. 2018, Macingo et al. 2019, Zheng et al. 2020, Sadeghi et al. 2021).

Metabarcoding studies of recent lake sediments have also produced mixed results: in Lake Baikal, Siberia, Yi et al. (2017) found the same four top taxa in both sediment and water samples – Chrysophyceae, Ciliophora, Metazoa and Cercozoa. However, in a study of 296

lakes across New Zealand, Dinophyceae emerged as the most abundant class in sediments, followed by Chlorophyceae (Pearman et al. 2023). Fungi, Ochrophyta and Ciliophora are also commonly reported in lake sediments (Capo et al. 2016, Wilden et al. 2021, Mitsi et al. 2023).

In rivers, studies sequencing communities from water samples found abundant Cryptophyta, Ochrophyta, Ciliophora and Fungi (Cruard et al. 2019, Cruard et al. 2020, Li et al. 2020, Lu et al. 2020, Xu et al. 2020, Yang et al. 2022), whilst sediment studies often highlight Ciliophora, Arthropoda, Ochrophyta and Bacillariophyceae (Xie et al. 2016, Hindshaw et al. 2017, Yang et al. 2022).

Diatoms (Bacillariophyta) have been one of the most extensively studied classes of algae and are well known as bioindicator organisms, included in many monitoring programmes under the Water Framework Directive. Zimmerman et al. (2014) found that metabarcoding retrieved over twice as many taxa as traditional light microscopy methods in river water samples, with a high degree of overlap between the two methods. Further Diatom metabarcoding studies are covered in section 3.

Research into aquatic fungi has not been nearly as extensive as those in other environments, such as soil, marine and air (Ruppert et al. 2019, Grossart et al. 2019). However, there have been some studies, particularly in lake ecosystems. For instance, Wurzbacher et al. (2016) found that fungal communities in lakes displayed significant turnover between littoral and pelagic habitats and between different substrates but noted that many taxa couldn't be identified to below order level due to the lack of completeness of reference databases. eDNA metabarcoding has shown that riverine fungal assemblages are structured by dendritic distance (Matsuoka et al. 2019). Metabarcoding has also been pivotal in detecting the pathogenic chytrid fungus (*Batrachochytrium dendrobatidis*) and sequencing the microbiome of amphibian species, critical for understanding and combating global amphibian population declines and extinctions (Kamoroff & Goldberg, 2017; Kueneman et al., 2017).

eDNA and metabarcoding in ponds

To date, eDNA metabarcoding of ponds has been used to study many of the taxa, and for many of the applications, listed above, but there are still large gaps in its application to the pond environment.

For example, much eDNA research in ponds and other small, lentic water bodies has focussed on the detection of amphibian species (Ficetola et al. 2008, Thomsen et al. 2012,

Goldberg et al. 2018, Eiler et al. 2018). In the UK, there has been a particular focus on the Great Crested Newt *Triturus cristatus* (Biggs et al. 2014, Rees et al. 2017, Buxton et al., 2018). Currently, a single-species eDNA assay is used to survey distribution across England. It comprises of publicly available records from over 5,000 ponds across England making it one of the largest eDNA survey in the world. Harper et al. (2018) found near-equivalence of *T. cristatus* detection from targeted PCR and metabarcoding methods, raising the potential for these samples to be metabarcoded with other primer pairs to survey other pond taxa (see Harper et al. 2019b).

Single-species and metabarcoding approaches in ponds have focused on detecting rare, endangered, commercially important, and invasive species. Examples include fish (for example, Turner et al. 2014, Doi et al. 2015, Robson et al. 2016, Evans et al. 2017, Li et al. 2018), reptile (Davy et al. 2015) and bird species (Ushio et al. 2018) for vertebrates, to crustacean and dragonfly species (Thomsen et al. 2012, Mauvisseau et al. 2018) for invertebrates. Studies of plant eDNA in ponds are rare, with the only example I could find being the detection of an endangered species (Newton et al. 2016.)

Like other freshwater environments and laboratory studies, metabarcoding of pond invertebrate communities using COI primers yields little overlap with morphological ID methods (Harper et al. 2021). An exciting potential use for eDNA ponds is to detect terrestrial or semi-aquatic species that use the pond or its catchment (Sales et al. 2019, Ishige et al., 2017), although so far, it has also shown little consistency with traditional visual survey methods.

eDNA is starting to be used to test ecological hypotheses. Harper et al. (2019b) used eDNA metabarcoding to determine of *T. cristatus* presence at a landscape level; for instance, there was a negative correlation with mammal species richness (as generated by metabarcoding). Ionescu et al. (2022) used 16S and 18S eDNA metabarcoding to examine the effect of land cover on pond communities in lowland Germany.

There are several challenges in using eDNA in pond environments and many unknowns (reviewed in Harper et al. 2019a). First, the origin of the eDNA: being small waterbodies, ponds have a high proportion of edge to core area, so a high proportion of eDNA may be transferred into the pond from the surrounding area. Birds, amphibians, waterfowl and herbivores may also transfer eDNA into ponds from further afield through feeding and bathing. Second, owing to low flow and lack of wind mixing, eDNA in ponds is likely to be patchily distributed both horizontally and vertically, and studies have found that eDNA

detection probabilities decrease dramatically within tens of metres of a source (Dunker et al. 2016).

Third, ponds vary significantly in their environmental conditions, such as temperature, nutrient levels, pH and light levels, which in turn will influence rates of eDNA degradation (Barnes & Turner 2016). Therefore, species detection via eDNA metabarcoding may be confounded by the successional state of the pond, its management (e.g., open canopy vs. closed canopy ponds), and surrounding land use. On the other hand, pond water and sediments can be highly anoxic (Sayer et al. 2013), which may slow eDNA degradation rates (Mejbel et al. 2022). Consequently, eDNA has been found to degrade within hours to days in water but persists for months to years in aquatic sediments (Barnes & Turner 2016, Turner et al., 2015).

Pond water can contain suspended solids, which make filtration difficult and lead to PCR inhibition (Harper et al., 2019). Finally, there are practical challenges in monitoring ponds that are common to all sampling methods: difficulty of access, changing water levels throughout the year, and deep sediments. However, some of the features that challenge eDNA metabarcoding in ponds also provide unique potential for this method in these environments. For instance, high interactivity with the surrounding landscape means that ponds could be “passive samplers” of terrestrial biodiversity. eDNA metabarcoding could also be used to investigate questions of connectivity and dispersal between ponds which have intrigued researchers for decades (see Section 4). The sheer numbers and varied conditions of ponds also make them excellent model systems (De Meester et al. 2005); for instance, comparing rates of eDNA degradation under different conditions.

Large research gaps exist in ponds, which create exciting prospects. Entire communities of bacteria, fungi, zooplankton, meiofauna, algae and eukaryotic micro-organisms have never been sequenced via metabarcoding from ponds. These taxa are the same as those that are understudied in ponds using traditional methods (see Sections 1 and 3). There is huge potential for research in this area; for instance, the role of pond microbes in biogeochemical cycles is unclear (Potvin et al., 2022), and microbial assemblages may be important bioindicators of pond ecological quality and anthropogenic impacts (see Sections 3 and Chapter 4).

Aquatic plants and other macrophytes have rarely been surveyed using eDNA metabarcoding in freshwater habitats, particularly in ponds. This is an area ripe for investigation, as macrophytes are key traditional freshwater bioindicators, yet require extensive time and expertise to survey (Pond Action, 2002). Areas of interest include rates of

plant eDNA shedding and persistence, detection overlap between eDNA and traditional methods (see Chapters 2 and 3), and ecological questions (dispersal, community assembly, see Section 4).

A new approach to eDNA metabarcoding in ponds is to move away from single-species detection towards characterising whole communities using multiple universal primers (Ionescu et al. 2022, Chapters 2, 3 and 4). This could provide a more complete picture of the biodiversity of these environments and reveal hitherto unknown ecological patterns and networks.

Section 3: Bioindicators in freshwaters, ponds and using eDNA.

Freshwater bioindicators: past and present

Bioindicators are biological processes, species, or communities that serve as tools to evaluate environmental quality and its change over time, with a particular emphasis on responses to human-induced pressures (Holt & Miller, 2010). They have been widely used since the 1960s and have been progressively developed to cover most habitat types worldwide. Bioindicators have the advantage of reflecting long term trends in environmental variables which spot samples, for example, of water chemistry measurements, may miss (Jones et al. 2016).

Simple bioindicators may be the presence of a single species or trait, or measures of community diversity and abundance of different taxa. However, more complex bioindicators are now standard use in freshwater environments. Biotic indices combine taxa abundances with sensitivity to different types of pollution to give a single biotic index or score for the sample site. Multimetric indices extend this by combining multiple measurements (e.g., taxa abundance, community diversity, functional feeding groups, abundance of pollution sensitive groups) into a single score (Li et al. 2010). Finally, multivariate approaches compare the sampled community to the "expected" community if no anthropogenic stress was present (Li et al. 2010, Ndatimana et al. 2023).

In rivers, the most common bioindicator groups are fish, benthic macroinvertebrates and periphyton (algae growing on surfaces) (Li et al. 2010), and in lakes these taxa are also

frequently used (Ndatimana et al. 2023, Liu & Stevenson 2017), with the addition of macrophytes and phytoplankton (Lyche-Solheim et al. 2013). Often multiple taxa are combined for the assessment of a site (e.g., phytoplankton, zooplankton, benthos and fish, US EPA 2019).

Benthic macroinvertebrates, particularly aquatic insects, are considered excellent bioindicators for freshwater environments as they are abundant, relatively easy to sample and identify by eye, and community composition is sensitive to changing physiochemical conditions in water (Li et al. 2010). Macroinvertebrates have been used as bioindicators of riverine ecosystems worldwide, forming part of governmental monitoring schemes in Western nations (e.g., the European Union, USA, Australia), and have been widely applied in the Global South (Eriksen et al. 2018). Macroinvertebrate indices commonly rely on identification to at least family level. UK rivers and lakes are monitored using macroinvertebrate assemblages, using a biotic index (the BMWP score) and a multivariate method (RIVPACS).

Diatoms (Bacillariophyta) are also common bioindicator organisms, and their use is recommended under the European Water Framework Directive, although different countries have developed varying approaches (Kelly et al. 2008). Sedimentary diatoms are also used as bioindicators of lake condition in the USA (US EPA 2017). Again, multiple measurements are often combined into a biotic index (Coste et al. 2008, Liu & Stevenson 2017) which require identification to family or species level using microscopy.

Fish, being both ecologically significant and economically and culturally important, serve as common bioindicators in freshwater environments. Their extended life cycles and diverse feeding strategies, along with varying pollution tolerances within this group, make them effective indicators of anthropogenic pressures (Li et al. 2010). Multimetric indices were developed in the 1980s (Karr et al. 1986) and multivariate approaches have been developed more recently (Joy & Death 2002).

Macrophytes (aquatic plants and algae growing in or near water) are more commonly used as bioindicators in lakes or wetlands than in rivers (Lyche-Solheim et al. 2013). Their use most often involves simple measures of abundance or diversity of different taxonomic groups or growth forms (such as in ecological classification of UK lakes, JNCC 2015), but sometimes more complex indices are used (e.g., Szoszkiewicz et al. 2020). There is a long history of using phytoplankton to assess trophic status of lakes, and most European Union countries monitor it regularly, although using different measurements or indices

(Pasztaleniec 2016). However, zooplankton is not included under the Water Framework Directive and so has received less attention in Europe (Pinto et al. 2023).

Microbes (bacteria, archaea and protists) are rarely used as bioindicators in legislative monitoring, save for those which have a direct impact on human health, and are cultivable (e.g., culturing *Escherischa coli* under the EU Bathing Waters Directive) (Sagova-Mareckova et al. 2021).

Freshwater bioindicators: future directions

The sampling, sorting and visual identification of specimens for conventional biomonitoring is costly, labour-intensive, destructive, and requires expertise and training. In the past decade, there has been considerable research effort into replacing these steps of freshwater biomonitoring with eDNA methods, which are thought to provide a greater amount of information with less time and effort (Stein et al. 2014). Consequently, many studies have been produced which compare the results of traditional, morphological identification methods with molecular-based methods for the same samples. Governmental regulators and scientists are actively working together to incorporate eDNA methods into legislative biomonitoring, for example STREAM in Canada (stream-dna.com) and DNAqua-net in Europe (Leese et al. 2016).

Studies of eDNA from freshwater macroinvertebrates have found varying degrees of overlap between the taxa identified from molecular and morphological methods, e.g., the proportion of morphologically identified taxa identified with eDNA varied between 24.4 % (Serrana et al. 2019), 67.5 % (Elbrecht et al. 2017a), and 73 % (Bush et al. 2019). There has been more success with community DNA from bulk samples (a sample containing whole organisms of study), or from preservative ethanol, and from sediment (Nichols et al. 2019, Hajibabeai et al. 2019). However, some methods still rely on hand sorting before sequencing, which would cancel out some the advantages of the molecular method in terms of speed (Carew et al. 2018). Although the taxonomic overlap is not perfect, studies have shown congruence between biotic indices calculated from molecular and morphological data for freshwater macroinvertebrates (Elbrecht et al. 2017, Brantschen et al. 2021).

The overlap between morphological and molecular detection methods for fish in freshwater environments is near 100 %, with eDNA often detecting additional species missed in traditional surveys (Olds et al. 2016, Hanfling et al. 2016, Shaw et al. 2016). Additionally, a recent meta-analysis found that 90 % of eDNA studies on fish found a positive relationship

between eDNA read count and fish abundance or biomass (Rourke et al. 2021). By contrast, low levels of taxonomic overlap have been reported for diatoms e.g., 13 % (Vasselon et al. 2017), 15.7 % (Rivera et al. 2017) and 28 % (Visco et al. 2015). However, despite these differences in species composition, the biotic indices and ecological quality assessments are often correlated for both approaches (Visco et al. 2015, Vasselon et al. 2017), although Rivera et al. (2018) found that they differed, most likely due to incomplete reference databases. There have been few studies comparing methods for plant communities, other than in lakes: Alsos et al. (2018) found a 30 % overlap in species detected by eDNA from lake sediments and visual surveys, whereas Drummond et al. (2021) found a 13 % overlap between eDNA from lake water and species lists.

Despite these successes, eDNA has not yet been adopted for wide scale legislative monitoring of freshwater ecosystems. There are several reasons for this: primer bias preferentially amplifies some taxa more than others, (Elbrecht & Leese 2017), reference databases are incomplete, (Pawlowski et al. 2018), and there is inconsistency between biotic indices used by different states and in different geographic areas (e.g., some rely on abundance of individuals), so an eDNA assay which provides the correct information in some areas may not be sufficient in others (Bush et a 2019).

Recently, there has been a movement towards alternative, innovative approaches. Instead of attempting to match morphological identification like-for-like, eDNA metabarcoding can be used to identify new bioindicators, such as bacteria, protists, fungi and micro- and meio-fauna, which are readily sequenced from environmental samples. Alternatively, taxonomic assignment is bypassed altogether, the so-called "taxonomy-free" approach. These sequenced community profiles can be compared and calibrated against known biotic indices or gradients of anthropogenic disturbance (reviewed in Cordier et al. 2020 and Pawlowski et al. 2018).

This approach has shown considerable promise, for instance in taxonomy-free approaches to assessing ecological quality of streams through diatom metabarcoding (Vasselon et al. 2017, Apothelez-Perret-Gentil et al. 2017), and assessing the impacts of salmon aquaculture via metabarcoding bacteria and ciliates in marine sediments (Cordier et al. 2018).

A clear future application of eDNA metabarcoding is the integration of microbes into routine freshwater biological monitoring. Microbes are sensitive bioindicators (Sagova-Mareckova et

al. 2021). Freshwaters are highly fluctuating ecosystems, and microbes respond rapidly to changes in the environment. Furthermore, they are intricately linked to many ecosystem functions, such as carbon, sulphur and nitrogen cycling. (Sagova-Mareckova et al. 2021, Cordier et al. 2020). Environmental genomic methods may be used to characterise entire microbial communities, or specific species or even genes which may indicate specific pollution (e.g., heavy metal resistance, Roosa et al. 2014).

eDNA Metabarcoding studies have successfully demonstrated the sensitivity of freshwater bacterial communities to mining waste (Staebe et al. 2018), agricultural pollutants in rivers (Xie et al. 2016, Chen et al. 2018), nano-pollution (Binh et al. 2014), pesticides (Pascault et al. 2014) and untreated wastewater (Martinez-Santos et al. 2018). High-throughput sequencing also has the potential to trace specific pathogens in water, such as those from faecal pollution (e.g., Vadde et al. 2019). Protist communities other than diatoms are less well-studied, but some taxa such as ciliates have shown responses to changing environmental conditions (Kulas et al. 2021). New biotic indices based upon bacteria taxa are being developed (e.g., Ji et al. 2019), and a recent study found a significant correlation between a taxonomy-free index based upon eDNA metabarcoding of both bacteria and microbial eukaryotes, and traditional water quality indices (Li et al. 2023). The next step for this area is to repeat the proof-of-concept studies over wider geographic areas.

All of this brings us back to the question of whether eDNA monitoring methods are in fact cheaper and faster than traditional methods. Estimates differ from the eDNA method being more expensive (Fernandez et al. 2018) to the two being roughly equivalent (Stein et al. 2014). Conversely, a recent meta-analysis concluded that molecular methods detect more species and were cheaper than morphological methods (Fediajevaite et al. 2021).

Pond monitoring past and present

Worldwide, there is a gap in the monitoring and protection of small waterbodies such as ponds, but also small lakes, headwater streams, springs and ditches (Biggs et al. 2017). In many areas, ponds are excluded from legal protection frameworks (Hill et al. 2018). For example, under the European Union Water Framework Directive (2000/60/EC; WFD) only lakes >50 ha and rivers with >10 km ^{2} catchment are monitored. In the UK, there are some protections for individual pond sites under the EU Habitats Directive and the UK Priority Habitats and Priority Species designations (Hill et al. 2018), but legislation does not require

regular nationwide monitoring and reporting on pond ecological condition. This may be set to change soon with the upcoming Natural Capital and Ecosystem Assessment (DEFRA 2022). However, as well as being impractical due to the sheer number of ponds (see above), inclusion in the WFD would not guarantee ponds' protection: in 2020, only 36 % of surveyed waterbodies in the UK and 16 % in England were in 'good' or better status, compared to the European average of 40 % (JNCC 2020, EEA 2021).

Despite this, nationwide pond monitoring in the UK has been carried out by charities, primarily the Freshwater Habitats Trust (formerly Pond Action), and scientific organisations since 1990. These have included a National Pond Survey of high-quality and impacted ponds in 1990-1993, the UK-wide Countryside Survey in 1996 and 2007, and resurveys of the high-quality ponds in 2016-2017 (Biggs et al. 2005, Freshwater Habitats Trust 2018). Together, these surveys provide a robust dataset of the changes in number and quality of ponds, with methods directly comparable to those used under the Water Framework Directive.

The monitoring of freshwater ponds has focussed on macrophytes, macroinvertebrates and amphibians. The standard method of surveying a pond in the UK is the PondNet method (described in detail in Biggs et al. 1998). This extensive method collects several different forms of data. This includes identification to species level of all macrophytes and macroinvertebrates, with macroinvertebrates sampled for four minutes in (ideally) three seasons. Additional data on pond water chemistry, physical attributes, macrophyte vegetation structure, the presence/absence of amphibians, waterfowl and fish, the surrounding landscape and pond management are also recorded.

As the method is lengthy and requires extensive expertise, many pond surveys have only collected data on macrophyte communities (e.g., Countryside Survey 2007, Williams et al. 2010). Recent data has supported the use of macrophytes as the primary bioindicator in ponds (but not lakes), finding that macrophyte morpho-diversity was an accurate predictor of macroinvertebrate diversity (Law et al. 2019). Data collected during the summer survey period can be used to calculate a multivariate index of pond quality (PSYM, comparable to RIVPACS, Pond Action 2002)

The Great Crested Newt (*T. cristatus*) is a protected species at the European and UK level (Conservation of Habitats and Species Regulations 2017) and as such, extensive monitoring effort have been directed towards this species, which lives primarily in ponds. In addition to the nationwide eDNA survey for GCN presence (see Sections 1 and 2, above), ponds are surveyed for their suitability for newts using the GCN "Habitat Suitability Index" (GCN HSI).

This index, developed by Oldham et al. (2000), is a composite index of ten different factors known to affect this species, such as pond area, macrophyte cover, shade, fish presence and water quality. The ten scores are combined to give an overall HSI score between 0 and 1, which can then be categorised into one of five categories from Poor to Excellent (ARG 2010, see Chapter 4 for more details). These scores have been calculated for thousands of ponds across England (Natural England 2019).

Pond monitoring: future directions

To ensure ponds' inclusion in the regulatory frameworks and nature-based solutions of the future, simpler, more rapid methods of monitoring need to be deployed, and eDNA metabarcoding is an obvious candidate. The ubiquity of ponds and their important biodiversity role in human-dominated landscapes also makes them prime candidates for monitoring via citizen science (Kelly-Quinn et al. 2022). The enormous potential for this kind of monitoring is demonstrated by a current citizen science programme which uses eDNA metabarcoding to examine taxonomic composition in urban ponds in England ("GenePools", Natural England 2023). Nonetheless, to integrate eDNA fully into pond monitoring, new methods must be scientifically robust, usable in ponds of different conditions and ideally, compatible with prior datasets.

Using eDNA analysis of ponds to discover new bioindicator taxa and/or to develop "taxonomy-free" methods of monitoring is an exciting prospect. Large-scale pond datasets, such as the HSI, PSYM and PondNet values, could be used to calibrate the communities sequenced from eDNA against known gradients of pond quality and create new biotic indices for ponds. This new sampling effort need not compromise current monitoring: Harper et al. (2019) found that detection of Great Crested Newt eDNA via eDNA metabarcoding was only slightly lower than with single-species qPCR. This opens up the prospect for whole pond communities to be surveyed alongside *T. cristatus* with a single sampling visit. However, the current protocol for GCN eDNA monitoring is a precipitation method, whereas studies have found filtration to be the optimum method for eDNA capture for metabarcoding (Deiner et al. 2015).

In summary, eDNA metabarcoding provides the opportunities to describe the microbial, algal, fungal and microfaunal communities of ponds which, until now, have been little studied and then only in a handful of sites, and provides an opportunity to survey pond communities year-round.

Section 4: Community assembly and eDNA

Humankind has drastically altered the face of planet Earth, with estimates of the proportion of altered land ranging between 75 % to 95 % (IPBES 2019, Kennedy et al. 2019). The oceans fare a little better, with 66 % human-impacted (IPBES 2019) whereas wetlands (both inland and coastal) have lost 86 % of their historic extent (IPBES 2019, Global Wetland Outlook 2018). The annual rate of loss of natural wetlands is calculated at –0.78 % a year (over three times the rate of loss of forests, Global Wetland Outlook 2018).

In the UK only 10 % of the historic extent of wetlands remains (Wetland Vision 2008). For ponds there has been an approximately 75 % reduction on historic extent and an accelerating loss in the past 100 years. Despite increased monitoring and advocacy efforts, conservation schemes and pond creation, ponds and freshwater habitats generally are undergoing alarming declines in biodiversity (see Section 1), and ponds in the UK are habitat for 10 % of “priority species” (Biodiversity Action Plan 1994).

How do we halt and reverse this decline? Do we need a pond in the corner of every farmed field or back garden? Or should we focus on preserving fewer, but larger, wetland areas (Freshwater Habitats Trust 2023)? Both combined is probably optimal for freshwater life, but not socially or economically feasible. This so called “land-sparing vs land sharing” debate is crucial in efforts to balance biodiversity conservation with other global needs, such as food production (Phalan et al. 2011). Globally, the trend is towards greater protection for large habitat patches over smaller habitat patches (Fahrig et al. 2017) e.g., the Lawton Review in the UK which recommended spaces for nature should be “more, bigger, better and joined” (Lawton et al. 2010). As we have seen previously, this strategy may not work for ponds. Many small ponds often have a greater biodiversity value than a single larger waterbody (see Section 1), and hydrological connectivity to other waterbodies may result in more pond pollution (Sayer 2014). To ascertain the threshold density of ponds in the landscape to maintain metacommunities, we need to gain a deeper understanding of the structuring of pond ecosystems. eDNA metabarcoding can play a pivotal role in providing answers to this important question.

Freshwater connectivity does appear to have an impact on pond biodiversity: greater species richness in ponds, and mitigation of species declines, has been linked to increased proximity to other waterbodies, location in semi-natural land and within large nature reserves. Proximity to urban areas or being located on the edge of reserves is linked with

increased declines (Oertli et al. 2002, Declerk et al. 2006, Williams et al. 2010, Williams et al. 2018). Hovarth et al. (2019) surveyed (temporary, saline) ponds in Austrian farmland. Between 1957 and 2010, 55 ponds had reduced to 30, and 17 zooplankton/rotifer species had been lost. This loss was far above that expected from loss of habitat alone (four species, calculated through species-accumulation curves).

However, it is difficult to disentangle connectivity from habitat quality: Hovarth et al. (2019) also found an increase in pond salinity in the same period, whereas the other studies mentioned above found significant effects of other variables linked to habitat quality, such as shade, shoreline development, water quality and grazing.

Connectivity is a broad term which can be applied to a range of, often quite different, concepts. In this section, I shall examine two broad approaches to investigating connectivity: habitat fragmentation and landscape connectivity, and their application to lentic freshwater environments. I shall look at two approaches to researching community assembly: bottom-up and top-down, or landscape scale, approaches. Finally, I will discuss the prospects for eDNA metabarcoding to examine hypotheses at the landscape scale.

Landscape connectivity

One conception of connectivity is that a connected landscape is one which is more “joined up”, with fewer, larger patches of habitat, compared to a fragmented landscape which has more, but smaller, patches of habitat. Structural connectivity is simply a description of the arrangement of habitat patches in a landscape and the linkages between them, such as corridors or inter-patch distances (Taylor, Fahrig & With 2006). It can be measured with several approaches which require the use of GIS or other landscape analytical tools (reviewed in With 2019, Chapter 5).

This conception owes much to the Theory of Island Biogeography (MacArthur and Wilson 1967). This seminal theory proposes that larger islands support more species than smaller islands due to lower extinction rates, and islands closer to the mainland can support more species than islands further away due to higher colonisation rates. Elegantly simple, the predictions were found to hold true empirically (Simberloff & Wilson 1970). The original authors suggested that the theory could be applied to human-fragmented landscapes, and in the ensuing decades fragmentation studies have often used patch size and isolation as a proxy for fragmentation (With 2019).

The effects of habitat loss and fragmentation on biodiversity have been extensively studied and reviewed. The species-area relationship (species richness increases as the habitat area

increases) has been found to hold true in terrestrial, freshwater, and marine ecosystems across the globe, and for organisms ranging from microbes to plants to mammals (Dakare et al. 2005). Habitat loss is the single largest driver of species extinctions and declines in abundance in terrestrial and freshwater ecosystems worldwide (IPBES 2019). Populations of species are predicted to experience a lagged response to habitat loss, and the full effects may not be seen for decades to centuries, termed “extinction debt” (Tilman 1994, Halley et al. 2016).

Whilst the effects of habitat loss are established, the effects of fragmentation on biodiversity are more equivocal. This is partly because habitat fragmentation and loss are often confounded, and many studies measure fragmentation at an individual patch level, when it should be measured at the landscape scale (Fahrig 2003). Fahrig (2017) states that fragmentation should be considered as a feature of a landscape rather than a process, that “for a given amount of habitat, a more fragmented landscape has more, smaller habitat patches and contains a greater total length of habitat edge”, calling this definition “fragmentation per se” (Fahrig 2017).

In ponds and many other habitats, current evidence does not confirm that a patchier or more fragmented landscape has a detrimental impact on species richness; instead, several smaller ponds hold a greater species richness than the equivalent area of a single large pond or lake (see Section 1).

There are other limitations to this approach. Fragmentation is often conceived of as a human-induced process, rather than a natural state of ecosystems. Additionally, the fragmentation paradigm, being based on the Island Theory, assumes that the intervening matrix between habitat patches is totally inhospitable for the target species. This is obviously untrue for many pond species, such as amphibians and Odonata, which feed in the surrounding terrestrial landscape and breed in ponds. The permeability of the matrix will alter drastically between wet and dry seasons and years, as may the size and arrangement of habitat “islands”.

Alternative approaches expand the definition of connectivity. Whilst landscape structural connectivity or fragmentation only accounts for the size and arrangement of habitat patches. landscape functional connectivity incorporates the response of an organism to a landscape. It is defined as “the degree to which the landscape facilitates or impedes movement among resource patches” (Taylor et al. 1993, Taylor et al. 2006, With 2019). Whereas structural connectivity remains the same across a landscape no matter what taxa are being studied, functional connectivity differs depending on the taxon (Taylor et al. 2006).

Aquatic connectivity can be expanded further to include flows of matter and energy, such as nutrients, carbon, or sediments between ecosystems (Jones et al. 2019). The transfer of matter and energy via water is referred to specifically as “hydrological connectivity” (Ormerod et al. 2011). Biological connectivity may occur with or without hydrological connectivity: compare a river flood event depositing sediment and allowing a fish species to colonise a pond, with the dispersal of a water beetle species via the air from one pond to another. Connectivity may even be extended temporally as well as spatially (e.g., tolerance of communities to drying events), and between land, sediment, surface water and groundwater as well as between surface water bodies (Ormerod et al. 2011, Mushet et al. 2019).

As with other concepts in freshwater research, aquatic connectivity has been less frequently applied to ponds than other waterbodies. Below I shall outline two major methods of studying biotic connectivity between ponds, and research to date; the direct dispersal approach and the metacommunities approach; and a brief overview of other methods.

Bottom-up approaches: dispersal, genetics, and occupancy modelling

The ability of aquatic species to locate and colonise new waterbodies within days of their creation has astonished naturalists for over 150 years (Darwin 1859, Darwin 1882, Talling 1951, McGuire 1963). A year after the creation of a new pond complex at Pinkhill Meadow, Oxfordshire, 37 species of wetland plant and 57 species of macroinvertebrates had colonised the four monitored ponds, and after seven years, the ponds held 20 % of all UK wetland plant and macroinvertebrate species (Biggs & Williams 2024, Williams et al. 2007).

Direct dispersal studies are one method of estimating the biological connectivity of a landscape. Dispersal refers to the movement of organisms or propagules from a source location (birth or breeding site) to another location where establishment and reproduction may occur (Nathan & Shohami 2016). Dispersal movements have potential consequences for gene flow between populations, whereas smaller movements within the home range (diurnal, stochastic or itinerant movements) do not.

Dispersal mechanisms for freshwater organisms may be active, facilitated by the organism itself, or passive, facilitated by some vector (for reviews of freshwater dispersal, see Bilton et al. 2001, Ormerod et al. 2011, Incagnone et al. 2015, Mushet et al. 2019). Active dispersal may be via water, overland, or via the air. Passive dispersal includes dispersal by animal, water (surface and potentially groundwater), wind, and sediment vectors. Many species can produce seeds, propagules or resting stages that can survive desiccation, sometimes for

tens to hundreds of years enabling re-colonisation of new water bodies from a sediment 'bank'. However, there are still entire groups of organisms for which dispersal mechanisms are unknown (Heino et al. 2015).

See the table below for examples of taxa which employ each dispersal mechanism.

Table 1.2: Modes of dispersal between freshwater bodies and examples of taxa which deploy them. Adapted from Mushet et al. (2019).

Type of dispersal	**Example taxa**
Active, water	Fish; lotic insects and crustaceans; some mammals, amphibians, and birds
Active, overland	Many amphibian, reptile and mammal species, some flightless insects, and other invertebrates (e.g., beetles, molluscs), some crustaceans
Active, air	Many aquatic insects that have a terrestrial adult stage; some fully aquatic diving beetles (Dytiscidae), water boatmen (Corixidae) and backswimmers (Notonectidae)
Passive, water	Most riverine plants, some freshwater molluscs and crustaceans, many others
Passive, overland (inc. animal vectors)	Microalgae, zoochorous plants, many invertebrates e.g., molluscs, leeches, micro-crustaceans, bryozoans
Passive, air	Microbes, anemochorous plants, weakly flying small insects, some zooplankton
Passive, sediment	Bacteria, fungi

Dispersal mechanisms and distances have been deciphered through a variety of methods, for instance: observations of eggs or seeds attached to larger organisms (e.g., Vanschoenwinkel et al. 2008), tests of viability of seeds or propagules after passage through digestive tracts (e.g., Darwin 1859), capture of aquatic insects in traps (e.g., Didham et al. 2012) and anecdotal evidence (for a more thorough review, see Figure 1 and Table 1 in Bilton et al. 2001).

The quantities of dispersing organisms or propagules at any one time may be vast: Soon et al. (2016) estimate that 500 million viable seeds are dispersed by dabbling ducks daily in Europe, whereas Csabai et al. (2006) recorded over 45,000 individual aquatic insect species attracted to black plastic sheets over six summer months (many actively dispersing aquatic insect species can detect polarised light indicative of a reflective water surface).

A widely accepted hypothesis posits that dispersal ability in passive organisms is negatively correlated with body size, whereas in active organisms, it is positively correlated with body size (Jenkins et al. 2007). Their extensive review of over 700 studies confirmed this relationship for active dispersers, but data for passive dispersers, especially small ones like microbes, was lacking. The question of whether microbial species have inherent dispersal limits remains a subject of debate (Nemergut et al. 2013, Van der Gast 2015). De Bie et al. (2012) provided supportive evidence for this hypothesis, as discussed in the following section.

The dispersal method influences species persistence. Research by Ozinga et al. (2009) revealed that plant species relying on water or mammal fur for dispersal, or lacking a persistent seedbank had a higher proportion of declining species compared to those relying on wind, birds, or having a persistent seedbank. This suggests that the loss of freshwater habitat, hydrological connectivity and wild megafauna in today's landscapes are contributing to species declines. Moreover, some species may experience declines only after their seedbanks are depleted.

Population genetics provides an indirect method of studying dispersal by quantifying gene flow between different populations. This method works via sequencing certain genetic markers in different individuals and populations across the study area and estimating genetic similarity and difference within and between populations. It’s assumed that high genetic divergence between populations is a result of low gene flow due to low dispersal, although differences may also be the result of selection and drift. Additionally, different rates on mutation between genetic markers will give different estimates of gene flow (With 2019 chapter 9).

Population genetic methods have often revealed strong genetic differentiation between relatively geographically close populations, even for small aquatic organisms with high rates of dispersal such as zooplankton and phytoplankton (Haileselaise et al. 2017, Rengefors et al. 2017). De Meester et al. (2002) reported evidence of high dispersal rates but low gene flow for passively dispersing organisms inhabiting lakes and ponds and hypothesised that this is due to strong founder effects (the “Monopolisation Hypothesis”). A combination of

landscape connectivity measures with genetic population methods have revealed isolation by landscape resistance rather than simply distance: for example, Ruggeri et al. (2019) found genetic structuring of bryozoan populations based on hydrological connectivity.

A third “bottom-up” approach is occupancy analysis, which compares potential species occupancy of sites in a landscape with actual occupancy (MacKenzie 2006). This method has been used to assess the effects of functional connectivity on amphibians (e.g., Cosentino et al. 2014, Swatasky et al. 2019). Niggebrugge et al. (2007) applied this approach to 20 gastropod species inhabiting marshes in the South-East of the UK, finding that all species only occupied a proportion of their potential habitat (21 – 87 %), and species occupancy decreased significantly as distance to nearest suitable habitat increased.

Top-down approaches: metacommunities approach

Determining the connectivity of habitats by quantifying the dispersal abilities of each individual species in that habitat, and how they interact with the landscape, is an arduous task. Besides, the distribution of species in patches of habitat is not only determined by dispersal at the landscape scale, but also by environmental constraints, biotic interactions and priority effects at the local scale, and speciation and extinction at broader scales.

Extending the previous metapopulation model of Hanksi (1998) and others, Leibold et al. (2004) formulated the metacommunity concept, which has since had a substantial effect on ecology. They defined a metacommunity as “a set of local communities that are linked by dispersal of multiple potentially interacting species” (Leibold et al. 2004). In this model, the composition of a community in a habitat patch (a pond, for instance) is dependent on the environmental conditions of the patch, the dispersal/colonisation ability of species, and the competitive ability of species, in varying amounts depending on the model used.

Leibold et al. (2004) put forward four models of metacommunity assembly, although it has since been recognised that not all metacommunities will adhere to one of these types (Brown et al. 2017). These are: neutral model, species sorting, patch dynamics and mass effects. In the neutral model, community structure is solely determined by random immigration, emigration, speciation and extinction. Species sorting assumes there is no dispersal limitation, and species’ presence is determined by suitable environmental conditions. Patch dynamics focusses on colonisation vs competition abilities of species, with colonisers dominating in isolated locations, and mass effects assumes dispersal abilities and environmental conditions are both significant.

Several researchers have suggested simplifying these four models to only focus on the relative effects of environmental filtering and dispersal limitation on communities, as these are the two fundamental processes at play (e.g., Heino et al. 2015). Usually, metacommunity studies examine whether there is a distance-decay pattern in community similarity along a spatial or environmental gradient, or both. This is typically computed using variance partitioning (see Brown et al. 2017 for more detail), or novel methods such as metacommunity assembly models (MAMs) (Brown et al. 2018).

One potential reason why meta-analyses of freshwater metacommunities have arrived at few generalisations, is because the studies compared are often over wildly different spatial extents (De Bie et al. 2012, Heino et al. 2015). For instance, Van de Meutter et al. (2007) studied macroinvertebrate communities in farmland ponds in Belgium over 300 ha and found no environmental or spatial structuring for active dispersers, and very limited evidence of spatial structuring for passive dispersers. By contrast, Fuentes-Rodrigues et al. (2013) examined macroinvertebrate communities of farmland ponds in southern Spain over 90,000 km ^{2}, and found environmental variation alone explained 4 % of the variation between communities, spatial variation around 15 %, and both variables also around 4 %.

Heino et al. (2015) in their review of freshwater metacommunities hypothesised a U-shaped relationship between spatial extent and spatial control of metacommunities: at very small extents, high dispersal will override environmental tracking. At large extents, dispersal limitation will inhibit species from tracking environmental changes. Intermediate extents are where the most environmental control (species sorting) is evident (Heino et al. 2015). Caution should be exercised over covering a very large extent as species may be structured phylogenetically, reflecting evolutionary events such as speciation.

Their second hypothesis is that the spatial control of metacommunities is expected to be greater for species with poorer dispersal ability, and vice versa. An important contribution comes from De Bie et al. (2012), who studied 99 Belgian farmland ponds over 30,500 km ^{2}. They surveyed these ponds for 12 different groups of varying body size, from bacteria to fish. Passive dispersers (bacteria, phytoplankton, rotifers, cladocerans, diatoms, macrophytes and molluscs) showed a clear relationship between increased body size and spatial limitation ($r=0.85$, $p=0.013$). By contrast, active dispersers by air (coleopterans, heteropterans and chironomids) showed lower spatial control and higher environmental control, than passive dispersers of the same size. Active dispersers overland or water (fish and amphibians) displayed high spatial control.

Soininen et al. (2011) found a similar result, with spatial structuring of communities across 100 Finnish lakes stronger for zooplankton than for phytoplankton and bacteria, and stronger at larger spatial scales than small. However, not all studies support these hypotheses. Garcia-Girón et al. (2019) used the MAM method to study community assembly of macrophyte species across 51 ponds in Northern Iberia, finding that wind-dispersed species showed a stronger signal of spatial limitation than other dispersal modes. Tornero et al. (2018) surveyed active and passively dispersing invertebrates and macrophytes across two different Mediterranean pondscapes, one of smaller (1.4 km) and one of larger extent (5.3 km), finding environmental control was significant for all three groups in both networks, but spatial control was only significant for active dispersers in the small network.

Whereas community assembly models for macro-organisms have been thoroughly studied, community assembly processes for micro-organisms are far less well understood. Throughout the twentieth century, it was commonly assumed that microbes had unlimited dispersal ability and were only 'selected' in a habitat by environmental factors (O'Malley 2008). In more recent times there has been growing evidence that microbial communities can also be spatially structured and follow similar community assembly rules to macro-organisms (Green and Bohannan 2006, Soininen 2012, Hansen et al. 2012). Much more research needs to be done to understand this area, particularly looking at different spatial extents and grains (Ladau & Eloe-Fadrosh 2019).

Therefore, the relative importance of environmental filtering and dispersal limitation on community assembly depends on the size of the organism studied, the dispersal method and ability of the organism, the spatial extent, and the features of the habitat itself.

Future approaches

eDNA metabarcoding provides a fantastic opportunity to study dispersal, reducing the labour of traditional mark-recapture studies, and clarifying dispersal distances for understudied biota. For instance, a time series sampling of eDNA in a newly created pond, coupled with sampling of freshwater habitats in the surrounding landscape, could elucidate dispersal distances, times, and origins of colonising organisms. This approach has been used to study the dispersal of single invasive species (e.g., Vimercati et al. 2018), and the colonisation of entire communities following glacier retreat (Rosero et al. 2021) but has yet to be applied to ponds to my knowledge. eDNA also has shown promise as a population genetics tool (Adams et al. 2019), both in aquaria and mesocosms (Andres et al. 2021, Marshall et al. 2019) and in wild aquatic habitats (Sigsgaard et al. 2017).

eDNA metabarcoding is already being used to investigate species occupancy at a landscape scale: Harper et al. (2019b) used 12S rRNA gene eDNA metabarcoding to sequence vertebrate communities across over 500 UK ponds. They found that GCN presence or absence was associated with the presence of other vertebrate species, and with abiotic variables.

eDNA metabarcoding also has a great potential to be used in landscape-scale community assembly models, particularly using multiple, universal primers to sample taxa of different body sizes (micro-organisms to macro-organisms) in the same environment and over the same spatial extent, permitting comparisons in community assembly rules between different taxa. For a recent example of this approach applied for coastal eukaryotic plankton, see Yan et al. 2023, and for ponds, see Chapter 4.

Another exciting prospect is combining the landscape connectivity, dispersal, and metacommunities approaches for ponds. Boothby (1997) used the term “pondscape” to delineate a network of ponds and the intervening terrestrial matrix, using GIS to estimate density and arrangement of ponds, and character of the matrix. Mushet et al. (2019) have recently expanded this concept to all freshwater bodies in a landscape with their concept of a “Freshwater Ecosystem Mosacic” (FEM), defined as “a collection of aquatic and wetland habitats in an inland landscape, and their occurrence within a terrestrial matrix”.

They hypothesise that a FEM landscape acts as a ‘filter’ on the species present, depending on dispersal mechanism. For instance, FEMs with a high density of lotic and lentic waterbodies, wet climate and semi-natural matrix are predicted to support species with the broadest range of dispersal traits, whereas decreasing connectivity progressively filters out species. For instance, FEMs with a low number of lotic bodies may not support water dispersed species, or FEMs in a dry climate may filter out overland dispersing species. Similarly, FEMs with a low density of waterbodies may not support short-distance dispersing species.

A potential future study could compare two pondscapes or FEMs of similar spatial extents, but with different characteristics (e.g., upland vs lowland, tropical vs temperate, high density vs low density), and use eDNA metabarcoding to sample either a single taxon or a range of taxa to compare community structuring and assembly.

Conclusion

Freshwater biodiversity loss is an accelerating global crisis. Ponds are relatively overlooked aspects of the freshwater environment, with little specific monitoring or protection, yet available evidence suggests their biodiversity importance outweighs their small size. They may be particularly crucial habitats for freshwater diversity in highly human-impacted landscapes such as agricultural and urban environments, and as such have the potential to be “nature-based solutions” to freshwater biodiversity loss.

Environmental DNA (eDNA) metabarcoding is a cutting-edge approach that has gained immense prominence in the scientific community and the broader media landscape in recent years. It has been applied to detect single species, entire communities, and even genomes and functional genes of those communities. Its non-invasive and non-destructive sampling methods, coupled with speed and scalability, hold the potential to redefine our understanding of the natural world. New applications for eDNA metabarcoding are emerging all the time.

In the context of ponds, eDNA metabarcoding has been used to detect a wide range of taxa, including amphibians, fish, reptiles, birds, and invertebrates. Nevertheless, using this method in ponds is not without its challenges. Factors such as eDNA origin, distribution patterns, environmental conditions, and practical obstacles pose difficulties that researchers must address.

Research gaps in eDNA metabarcoding of aquatic plants, and whole-community characterisation offer opportunities for exciting discoveries in ponds. eDNA metabarcoding will allow us to delve into the microbial, algal, fungal, and microfaunal communities of ponds, offering new insights into these understudied groups. As we look ahead, the continued development and application of eDNA metabarcoding in pond environments have the potential to expand our understanding of these intricate and biodiverse ecosystems, contributing to advancements in ecological research and conservation efforts.

The use of bioindicators in freshwater environments, including rivers, lakes, and ponds, has played a crucial role in monitoring environmental quality, and assessing human-induced pressures over the years. Bioindicators, ranging from simple species presence to complex

biotic indices, have provided valuable insights into the health of these ecosystems, and are used widely to guide regulatory and conservation efforts.

Traditionally, these assessments have relied on labour-intensive and time-consuming morphological identification methods. However, recent developments in eDNA techniques are changing the landscape of freshwater biomonitoring. While eDNA methods show promise in terms of rapid data collection and the potential to expand the range of bioindicator organisms and deploy “taxonomy-free” methods, there are challenges to overcome, such as primer bias, incomplete reference databases, and variations in biotic indices. Nevertheless, the field is evolving, and researchers are actively working on refining these techniques.

In the context of pond monitoring, a significant gap exists in the regulatory frameworks for smaller water bodies, including ponds. Nevertheless, organisations and scientists have contributed valuable data on pond ecosystems over the years, focusing on macrophytes, macroinvertebrates, and amphibians. These data have been critical in understanding pond health and changes in biodiversity. Looking ahead, the inclusion of ponds in regulatory frameworks and the development of more efficient monitoring methods, such as eDNA metabarcoding, are essential. The potential for citizen science involvement in pond monitoring is a promising avenue to explore.

The large reduction in pond numbers and density in the last century and ongoing declines in wetland area is predicted to cause species extinctions over and above that caused by habitat loss alone, due to loss of connectivity prohibiting successful colonisation events. However, the effects of habitat loss and connectivity are difficult to disentangle. When studying ponds, accounting only for distances between ponds (structural connectivity) is insufficient as many pond species also inhabit the surrounding landscape, and can move between different freshwater environments.

The effects of pond connectivity on pond biodiversity can be studied via elucidating dispersal mechanisms and distances of different taxa, by population genetics and by occupancy modelling. One drawback of these studies is that they can only focus on certain taxa at a time. Top-down approaches, such as comparing the relative contributions of environmental and spatial drivers to community assembly, have provided empirical support to models of the interaction between body size and dispersal method and community assembly. Some evidence is contradictory, and studies are difficult to compare as they have been conducted over different spatial extents and grains.

Moving forward, eDNA metabarcoding will be a key tool for advancing our understanding of dispersal, population genetics, and testing landscape-scale hypotheses. Furthermore, the integration of landscape connectivity, dispersal mechanisms, and metacommunities in the study of "pondscapes" or "Freshwater Ecosystem Mosaics" offers an exciting avenue for future research.

Chapter 2: From microbes to macrophytes: optimising multi-taxa eDNA metabarcoding methods in freshwater ponds

Abstract

This chapter reports on a proof-of-concept approach to whole-community environmental DNA (eDNA) metabarcoding of freshwater ponds. Twenty ponds (lentic freshwater bodies <5 ha and <5 m deep) in Pinkhill Meadow, Oxfordshire, UK, were surveyed on a single day in June 2020. Sediment and water were sampled from the ponds, from which eDNA was subsequently extracted, amplified and sequenced using an Illumina MiSeq. Each sample was amplified with five different primer pairs to survey different components of the pond community: prokaryotes (16S rRNA gene), eukaryotes (18S rRNA gene), animals (COI mtDNA), fungi (ITS2 gene region) and green plants and algae (ITS2 gene region).

After taxonomic assignment and filtering by confidence, the eDNA survey detected 887 taxa identified to genus or species level across the entire site, 4.9 times the number identified at the most recent traditional survey conducted in 2016. There was very low overlap between the two datasets (10 % of macrophyte species and 3 % of macroinvertebrate species), however eDNA methods detected many taxonomic groups for which data on abundance and distribution in pond environments is lacking, such as bacteria, meiobenthos, zooplankton, phytoplankton and algae.

There was evidence of migration of eDNA from external environments into the ponds, with 70 % of macrophyte reads in water samples identified as terrestrial in origin. Bioinformatic processing approach had little effect on conclusions drawn from community composition analyses. Whether data was unfiltered (taxonomy-free), read abundance weighted or

presence-absence, sample type (water or sediment) and pond type both had a significant structuring effect. Combining multiple water samples from a single large (0.5ha) pond on one filter detected fewer taxa than if samples were filtered separately. Overall, this study provides important considerations for the design of future pond eDNA metabarcoding research.

Introduction

Environmental DNA (eDNA) metabarcoding is an emerging method for monitoring biodiversity. In brief, it involves taking an environmental sample (for instance water, soil, sediment, biofilms, or even filtered air), extracting the DNA contained within, and amplifying this eDNA via a polymerase chain reaction (PCR) using primers, which may be highly specific or general in the taxa they amplify. The resulting copies of the DNA sequences are then sequenced using high throughput sequencing (HTS)), and the millions of DNA 'reads' that result can be used to infer the biodiversity of the original environmental samples and hence, the original habitat (Taberlet et al. 2018).

Despite its relative novelty, eDNA metabarcoding methods have now been used around the globe, in habitats as diverse as the deep ocean, tropical rainforest, urban streets and arctic streams (Ruppert et al. 2019, Littlefair et al. 2023), although coverage is still lacking in many areas, particularly in the tropics (McGee et al. 2019). However, as an emerging field, eDNA metabarcoding methods are still in flux, and the method must be optimised for different ecological habitats and applications (Ruppert et al. 2019), particularly if it is to become the standard biomonitoring tool many claim it will be (Schenekar 2023).

eDNA metabarcoding has been widely applied for biodiversity monitoring in freshwater ecosystems worldwide (for reviews, see Schenekar 2023, Belle et al. 2019). Freshwater ponds have been a relatively common subject of eDNA metabarcoding studies (reviewed in Harper et al. 2019a). Most eDNA applications in ponds to date have focussed on monitoring a single species using a targeted qPCR assay. These species are often large, protected or charismatic amphibians such as the American bullfrog (*Rana catesbeiana,* Ficetola et al. 2008) the Great Crested Newt (*Triturus cristatus,* Biggs et al. 2015) or the Natterjack Toad (*Epidalea calamita* (Reyne et al. 2021), or invasive aquatic plants (*Egeria densa,* Fujiwara et al. 2016) and animals (crayfish *Procambarus virginalis,* Mauvisseau et al. 2018). When eDNA metabarcoding has been deployed to look at whole communities, these have most

often used a single primer pair and targeted vertebrates (e.g., mammals: Harper et al. 2019c, fishes: Di Muri et al. 2020, amphibians: Moss et al. 2022) or macro-invertebrates (Harper et al. 2021). Compared to rivers or lakes, fewer pond studies have used universal primers or multiple taxonomic markers to detect a wider taxonomic breadth of organisms (e.g., 18S rRNA gene to detect Eukaryota, 16S rRNA gene to detect Bacteria and Archaea, the Cytochrome Oxidase c subunit 1 gene to detect metazoans).

More recently assays targeting a wide range of taxa are being used in ponds, not only to monitor or assess biodiversity, but also to test ecological hypotheses. Macingo et al. (2019) used 18S rRNA gene eDNA metabarcoding to assess the diversity of unicellular eukaryotes across mountainous freshwater pools in Greece (but did not distinguish between smaller ponds and larger lakes). Ionescu et al. (2022) sampled the eukaryotic and prokaryotic diversity of kettle hole ponds in Germany using 18S rRNA and 16S rRNA marker genes, with a view to distinguish the effects of different land use types on bacterial, microbial eukaryote and invertebrate pond communities.

Using eDNA metabarcoding for biomonitoring of freshwater ponds has promising prospects, for instance, increasing the total number of ponds monitored (Harper et al. 2019), detecting rare or cryptic taxa (Goldberg et al. 2018) and determining terrestrial species presence/absence in the area surrounding the pond (Ushio et al. 2017). Using universal primers and multiple taxonomic markers is an exciting new application for pond eDNA metabarcoding, as it will increase the taxonomic coverage of monitoring far beyond the macrophytes (large plants and algae such as stoneworts) and macroinvertebrates used in traditional visual-based monitoring methods (e.g., PSYM, Pond Action 2002). Groups of organisms like Bacteria, Fungi, microbial eukaryotes, algae, phytoplankton and zooplankton have rarely been monitored in ponds, and very little is known about their ecology or distribution in these environments.

There are still challenges and unknowns in developing eDNA metabarcoding in ponds (reviewed in Harper et al. 2019a). The lack of flow and water mixing in ponds means that eDNA is likely to be patchily distributed, and local pond conditions may cause differences in rates of eDNA degradation. Other challenges include the seasonal drying experienced by many ponds and methodological difficulties caused by turbid water (slow filtration and PCR inhibition).

Established metrics of pond and other freshwater ecosystem health rely on identifying a taxonomically narrow group of bioindicator organisms (such as macrophytes, freshwater macroinvertebrates or diatoms) to genus or species level, using visual-based identification

methods such as sweep-netting and microscopy (Mainstone et al. 2018, Pond Action 2002). It has been suggested that eDNA metabarcoding could replace the identification stage of these methods, and result in cost and time savings (Baird & Hajababei 2012, Thomsen & Willerslev 2015). However, this promise has yet to be realised, as studies comparing detection between traditional, visual identification-based freshwater macroinvertebrate monitoring and eDNA metabarcoding to date have found incomplete overlap in species detection, between 25 % and 80 % (Serrana et al. 2019, Bush et al. 2019). A recent study of 18 ponds comparing traditional sweep-netting for macroinvertebrates with eDNA metabarcoding of the COI gene found 0 % overlap between species detected (Harper et al. 2021). This mismatch is thought to be due to current COI primer pairs primarily amplifying non-target taxa (Leese et al. 2020). Few studies have used metabarcoding to characterise the macrophyte community: however, Alsos et al. (2018) found that 31 % of plant species found within 2 m of a lake shore were also detected using eDNA from the lake water, whereas Drummond et al. (2021) reported only 13.2 % of the species visually detected in a lake and catchment were also detected in eDNA from lake water.

One currently open question in eDNA metabarcoding methods is whether the abundance of reads of a particular taxon in a sample can be used as a proxy for the abundance of that taxon in the environment (Di Muri et al. 2020). There are some instances of this relationship holding, for example, many studies have found a positive relationship between fish abundance or biomass and eDNA read abundance (Rourke et al. 2021). However, many more have found a mismatch between read abundance and known species abundance, particularly in studies of macroinvertebrates (Elbrecht & Leese 2015, Garrido-Sanz et al. 2021).

These mismatches could be due to many reasons, such as differences in body size, genome size, PCR amplification efficiency or biases in the bioinformatic pipeline towards or against certain species (Luo et al. 2023). Corrections have been proposed, for instance, pre-sorting organisms by size (Elbrecht et al. 2017) or using spike-in DNA (these corrections are reviewed in Luo et al. 2023). Many studies assume that for smaller, more numerically abundant organisms such as bacteria and eukaryotic microbes, relative read abundance can be used as a proxy for relative taxon abundance in the environment (e.g., Ladin et al. 2021), but variable copy numbers of the 16S rRNA and 18S rRNA genes can bias these estimates (Gao & Wu 2023, Gong & Marchetti 2019). Some analyses circumvent the problem by using presence-absence data in analyses rather than read abundance, but when metrics based on

presence-absence and read abundance are compared, often the outcomes are different (Ionescu et al. 2022, Machler et al. 2021).

The overall aim of this study was to optimise eDNA metabarcoding methods in ponds, using twenty ponds in a small wetland nature reserve of approximately five hectares, Pinkhill Meadow in Oxfordshire, as my study site. These are some of the most well-studied pond ecosystems in the UK, with macrophyte and macroinvertebrate species presence/absence data spanning 30 years, and amongst the most biodiverse ponds in the UK (Freshwater Habitats Trust 2019, Williams 2017).

On a single day in June 2020, 42 water and 42 sediment samples were collected from across these ponds, from which eDNA was extracted, and then amplified with five separate primer pairs, targeting different sections of the pond community. These were bacteria and archaea (16S rRNA gene), eukaryotes (18S rRNA gene), fungi (fungal ITS2 gene region), animals (COI mtDNA) and green plants (plant ITS2 gene region). The resulting amplified eDNA was sequenced on an Illumina MiSeq and sequences assigned to genus or species level.

This study had five separate research questions relating to the methodology of eDNA metabarcoding in ponds.

1. Do the five primer pairs listed above amplify eDNA in pond water and sediments? (At the time this research was conducted, no prior research could be found using these assays in pond environments).
2. If amplification is successful, which of the primer pairs, along with reference databases provide the best balance between broad and deep taxonomic coverage and confident taxonomic identification?
3. In order to accurately sample the ecological community of a larger pond (0.5 ha) using eDNA metabarcoding, are multiple water samples from separate locations required, or is it sufficient to sequence a single merged water sample?
4. How do bioinformatic processing decisions (e.g., using presence/absence or read-abundance weighted data, filtering depending on confidence of taxonomic assignment) affect the resultant community composition metrics for the different ponds and communities?
5. Finally, do eDNA metabarcoding methods targeting macroinvertebrates and macrophytes have good taxonomic overlap with traditional, visual-based survey data?

Methods

Study site

Pinkhill Meadow nature reserve, Oxfordshire, is a wetland complex of over 40 ponds, with varying sizes, shapes, depths, permanence and water sources (Figure 2.1). The ponds were excavated in 1990 to provide a practical case study for developing ideas on pond and wetland creation and management (Freshwater Habitats Trust 2019). The ponds were monitored, using traditional, visual identification-based methods, for wetland plants and invertebrates annually in the first decade, and then at intervals up until the present day (Biggs & Williams 2024).

The site is around 4 ha in size, and located in a meander on the River Thames, adjacent to Farmoor reservoir, to the west of the city of Oxford. Around 1.5 ha of the site is unimproved floodplain meadow and is grazed by a small herd of cattle between October and March and the hay cut once a year. Other than this, there is no management of the site, and it is not accessible to the public.

The underlying geology of the site is a clay (alluvium) layer overlying a gravel aquifer (Williams 2017). The alluvium layer is thicker at the northern end of the site, and so the ponds in this region have a clay substrate and tend to be semi-permanent, drying out in drought years. The ponds at the southern end of the site are dug into the gravel aquifer, and so are groundwater fed. Twenty distinguishable water bodies were chosen in December 2019 for sampling for water, sediment and abiotic variables at six roughly bi-monthly sampling events in 2020 (for results from other sampling events, see Chapter 3).

Figure 2.1 shows a map of the site with locations of sample points named in red. Some larger waterbodies were sampled in multiple locations around the perimeter of the pond. Smaller waterbodies were sampled in a single location near the perimeter. For full details of water, sediment and blank samples per pond, see supplementary Table S2.4.

Full sampling of water and sediment over all 20 ponds was carried out over a single day: 2nd June 2020. In-field measuring of abiotic variables was performed the subsequent day. In total, 42 water and 42 sediment samples were collected from across the site. The subsequent day, water samples were collected from 15 ponds (six samples were collected

from the Main Pond, and a single sample from 14 other ponds), and analysed in a laboratory for additional water chemistry variables.

Depending on prior knowledge and site observations (Williams 2017, Williams et al. 2010, P. Williams personal communication Jan 2020), the 42 sampling sites were divided into one of four “pond types” depending on pond substrate, location on site and canopy cover. These were “Main Pond” (n=18, clay substrate, part of the large pond and scrape complex, open, 5000 m^2 in area), “Surface-water pond” (n = 5, clay substrate, open, two ponds ca. 150 m^2 each), “Experimental Pond” (n=7, clay substrate, medium canopy cover, seven ponds each 100 m^2) and “Gravel Pond” (n=12, gravel substrate, high canopy cover, nine ponds ranging from 20 m^2 to 400 m^2).

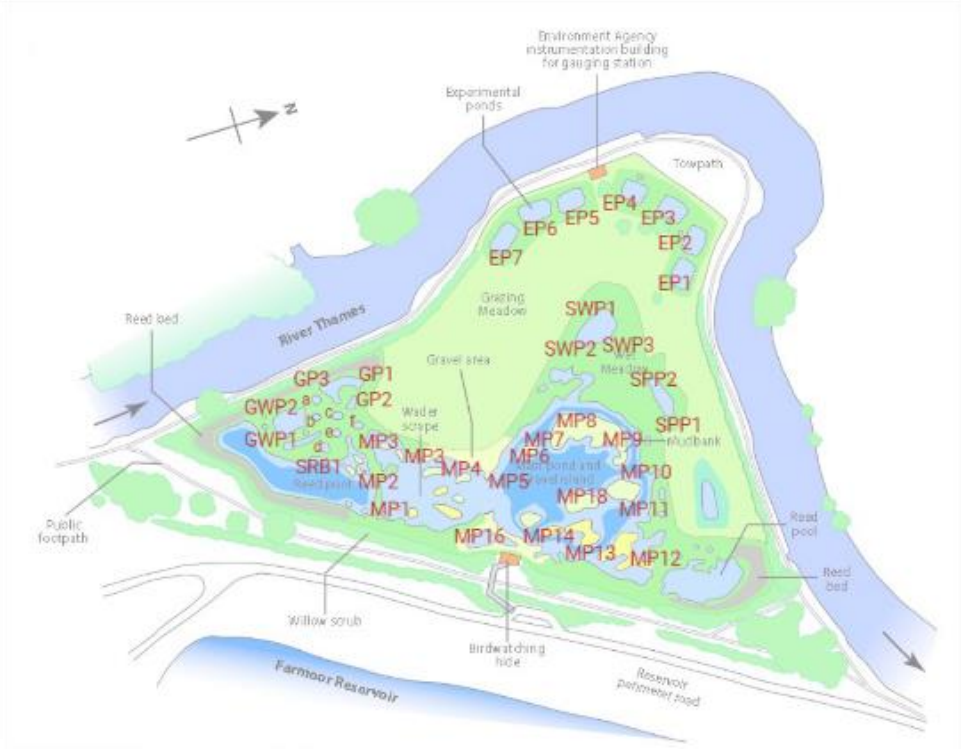


Figure 2.1: Pinkhill Meadows pond complex site, with key human and natural features labelled. All 42 sample points overlaid in red text. MP = Main Pond, SPP = Semi-permanent pond, SWP = Surface water pond, EP = Experimental pond, GWP = Groundwater pond, GP = Gravel pond, SRB = Southern reed bed.

Water and sediment sample collection

Water samples were collected using 1l plastic bottles which had been sterilised via a laboratory acid washer or soaking in 10 % bleach. Nitrile gloves were used to collect water

and changed between samples. Bottles were triple washed in pond water before a sample was taken. Water samples were taken around 10cm from the surface of the pond in all cases (a compromise between the tendency of eDNA to sink (Turner et al. 2015, Harrison et al. 2019) and avoiding disturbing the sediment). Water was transported in a cool box with ice packs to laboratory facilities within four hours of collection, where it was stored at 4°C for a maximum of 24h before filtration, as recommended by Curtis et al. (2021).

A water “field blank” was generated by filling a sterilised 1l plastic bottle with deionised water, transporting it to, around and from the field in the same method as the other sample bottles.

Sediment sample collection and storage

Sediment samples were collected from the same location as the water samples. This was both for the sake of practicality and to reduce the influence of external factors e.g. light levels on differences in community composition between sediment and water samples. Sediment samples were taken using a custom sediment sampler comprising of a tubular aluminium rod with a holder for a 50 ml centrifuge tube, both of which had been sterilised using an acid wash or 10 % bleach before use. The scoop was triple washed in pond water between each sample. Sediment samples were taken from the surface of the sediment, within the top ~10cm.

To attempt to prevent cross contamination between sample sites, the sediment scooper was wiped with 10 % bleach between uses and left to air dry. Sediment samples were transported to the laboratory within 4 hours of collection and stored at -20 °C.

Measuring abiotic variables

In the field, an ultrameter II (Myron L Company) was used to measure pH, temperature, Total Dissolved Solids (TDS), Conductivity and Oxidative-Reductive Potential (ORP) at each sample point. Measurements were taken at the same time as the water and sediment samples were collected. The sensor was triple washed in pond water before the measurements were taken. For each sample point, three measurements were taken, and the mean calculated.

Water depth at each sample point was measured using a metre rule. Water samples for water chemistry analysis were collected using a wide-mouthed bucket which was rinsed with sample water before sampling. From this larger sample, a smaller 100 ml sample was taken to quantify total suspended solids (TSS) in the laboratory. Two 60 ml samples were subsampled from the bucket and filtered using 0.45um filter on site. The bucket was then

stirred, and a third 60 ml sample was collected and not filtered. The two filtered samples were analysed for 1. total Soluble Reactive Phosphorus (SRP) and 2. chlorophyll, ammonia (NH_4), Total Dissolved Nitrogen (TDN) and Total Organic Carbon (TOC). The unfiltered sample was analysed for Total Phosphorus (TP). The water chemistry analysis was carried out using the methods described by Bowes et al. (2018).

CanopyApp (University of New Hampshire, ver. 1.0.4) was used to estimate tree canopy cover at each sample point in July 2020. The site was visually assessed during each sampling event and photographs taken to record which waterbodies were hydrologically connected or separated due to flooding or drying.

Filtration and eDNA extraction

Filtration was chosen over precipitation methods due to the evidence of a greater rate of eDNA recovery (Hinlo et al. 2017). An average of 1000 ml of water was collected for eDNA filtration at each sample point. For 14 smaller ponds, only one 1000 ml sample was taken. For three medium ponds (>100 m^2, GWP, SPP and GP1), two 1000 ml water samples were taken in the field. In the laboratory, each of these were sub--sampled to 500 ml which were combined to make one 1000 ml sample which was subsequently well mixed and filtered.

One medium-size pond (SWP) was sampled in a similar way, except three 1000 ml samples were taken in the field, and these were sub-sampled for 330 ml each which were combined to make one 990 ml sample. For the large main pond, up to 17 separate 1000 ml samples were collected in plastic bottles in the field. In the laboratory, ~60 ml of each of these samples were subsampled and combined to make a merged Main Pond sample of ~1000 ml, which was subsequently filtered. The ~940 ml of the 17 separate samples (MP1 – MP17) were also filtered separately.

All samples were filtered in a clean laboratory environment using vacuum pump filters sterilised via acid washing or soaking in 10 % bleach. Each 1000 ml sample was initially pre-filtered with a 12um Cellulose-Nitrate filter (Whatman, AE100) and subsequently filtered with a 0.45um Cellulose Nitrate filter (7141 114). The prefiltering was carried out to reduce the influence of large organic particles on downstream analysis, and the filter size chosen due to best available evidence (Li et al. 2018). The filter papers were removed from the filters using sterilised tweezers and placed in 5 ml centrifuge tubes, which were stored at -20°C until extraction.

DNA from water samples was extracted from the stored filter papers after defrosting using the standard protocol of the E.Z.N.A. water DNA extraction kit (Omega Biotek). A water extraction blank was produced by following the normal extraction protocol but omitting any sample. Concentration of DNA was determined using a Qubit Fluorometer (Invitrogen). DNA was stored in Elution buffer in 1.5 ml microcentrifuge tubes at –20'C. Total volume ranged from 70 µl – 100 µl.

DNA from sediment samples was extracted by defrosting the full sediment sample and then subsampling 0.25g using scales, a sampling boat and disinfected tweezers. DNA was then extracted using a DNeasy Powersoil kit (Qiagen). A sediment extraction blank was produced by following the usual protocol but omitting any sample. Concentration of DNA was determined using a Qubit Fluorometer (Invitrogen) or Nanodrop 8 sample (Thermo Fisher Scientific). DNA was stored in Elution buffer in 1.5 ml microcentrifuge tubes at –20'C. Total volume ranged from 80 µl – 85 µl.

All extraction was carried out using sterilised procedures e.g., gloves, and the use of 99 % ethanol and bleach to sterilise equipment and laboratory benches between any handling of DNA samples.

Amplification and sequencing

Initially 20 primer pairs were tested on six randomly chosen samples (three water and three sediment), along with an extraction blank and PCR blank, to ensure positive amplification. From these, five primer pairs were chosen, which amplify different DNA or RNA targets and produce different length fragments, in order to detect different sections of the pond ecological community (see Table 2.1 below).

Table 2.1: Information about the five primers used in the eDNA survey of Pinkhill Meadow ponds, June 2020.

Primer pair name and reference	**Fragment length**	**Gene amplified**	**Community primer optimised to target**
BF1/BR2 (Elbrecht & Leese 2017)	316bp	COI mtDNA	Freshwater macroinvertebrates
Euk1391f/EukBr (Amaral-Zettler et al. 2009 and Stoek et al. 2010)	45-260bp	18S rRNA	Eukaryotes

515F/806R (Caporaso et al. 2011 and Walters et al. 2015)	390bp	16s rRNA	Bacteria
ITS2-S2F/ITS4_R (Ihrmark et al. 2012)	300-460bp	ITS2 rDNA	Vascular plants
ITS7_F/ITS4_R (Ihrmark et al. 2012)	~250bp	ITS2 rDNA	Fungi

DNA was amplified using a two-step PCR approach. Firstly, all samples were amplified with a modified primer (amplicon primer with Illumina MiSeq sequencing primer and pre-adaptor added). 1-2 PCR blanks were also amplified at this stage (normal PCR reagents, but with molecular grade water added rather than any sample). Then Step 2 PCR was carried out to add on the barcodes (Illumina MiSeq index) and flow-cell adaptors. Steps 1 and 2 were repeated for the four other primer pairs. Therefore, each sample had five two-step PCRs carried out on it, one per primer pair. The PCR conditions are detailed in table S2.1.

Amplicons were normalised using the SequalPrep Normalisation Plate Kit, 96-well (Invitrogen, Carlsbad, CA), gel purified using the QIAquick gel extraction kit (Qiagen Group), and quantified using Qubit high sensitivity dsDNA Assay kit (Invitrogen, Carlsbad, CA). The resultant amplicon library was sequenced at a concentration of 9 pM with a 0.675 pM addition of an Illumina generated PhiX control library. Sequencing was performed on an Illumina MiSeq platform using MiSeq Reagent Kit v3 (Illumina Inc., San Diego, USA).

Bioinformatic analysis

For the 16S rRNA gene, 18S rRNA gene, COI mtDNA and Fungal ITS2 (fITS) sequences, raw reads were processed through the DADA2 pipeline ver. 1.8 (Callahan et al., 2016) in R (R Core Team, 2018). Briefly, adapters and primers were initially removed from the raw reads using cutadapt (Martin 2011), then amplicon reads were trimmed to maintain Q score > 30. This occurred at 250 and 200 bases for 16S 18S and COI sequences, and 250 and 220 bases for fITS sequences, (forward and reverse reads respectively.) The results filtered with DADA2 default settings, except for the maximum number of Ns (maxN) = 0 and maximum number of expected errors (maxEE) = c(5,5).

Sequences were dereplicated and the DADA2 core sequence variant inference algorithm applied. Processed forward and reverse sequences were merged using the mergePairs function, and a sequence table was constructed from the resultant, merged amplicon

sequence variants (ASVs). Chimeric sequences were removed from the ASV table using remove BimeraDenovo with default settings.

To assign taxonomy to 16S, 18S and fITS sequences, the online IDTAXA tool from the R package DECIPHER ver. 2.14 (Murali et al. 2018) was used. The reference databases used were: for 16S sequences, SILVA ver. 138.1 (10th March 2021); for 18S sequences, PR2, version 4.13 (17th March 2021), and for fITS sequences, UNITE ITS (ver. 2020, February 2020), all of which are provided within the DADA2 package. This function gives a taxonomic classification to genus level in the case of 16S and 18S, and to species level for fITS, and provides confidence scores for each level of classification. The confidence threshold for assignment in all three cases was set at 60 %. For the COI sequences, the Python package BOLDigger was used (Buchner & Leese, 2020), which matches sequences to both private and public sequences stored in the Barcode of Life Data systems (BOLD, Ratnasingham & Herbert, 2007). This program gives a taxonomic classification to species level and provides an overall confidence score.

For the plant ITS2 (pITS) sequences, a different process was followed. Raw reads were processed and taxonomically assigned using the HONEYPI pipeline implemented in Python 2.7 (Oliver et al. 2021). The HONEYPI pipeline removes adaptors and quality filters the raw reads using TrimGalore v.0.6.4, and then uses the DADA2 pipeline to generate an Amplicon Sequence Variant (ASV) abundance table containing chimera-removed, high-quality error-corrected sequences. For each ASV, conserved regions flanking ITS2 are removed with ITSx v.1.1b; and resulting sequences taxonomically classified using the naive Bayesian classifier against an in-house ITS2 database of 966,676 sequences (25th March 2020). Unless stated otherwise, default parameters were used for the steps listed. Each of the ASVs was taxonomically assigned to species level by the RDP classifier within the HONEYPI pipeline and given a confidence level of assignment between 0 and 1.

Using the R package ‘phyloseq’ (version 1.42.0, McMurdie & Holmes 2013), each of the five datasets went through the same three steps: removal of low-read samples, removal of ultra-rare taxa (less than 3 reads and only in 1 samples) and rarefaction to the median sampling depth. Removing low-read samples and taxa aids in the removal of false positive detections (Shirazi et al. 2021, García-Machado et al. 2023). Rarefaction is a widely used and statistically valid way to normalize sample size, and so control for the uneven sequencing effort between samples (Weiss et al. 2017, Schloss 2023). For the 16S, 18S, COI and pITS datasets, samples below 4,000 reads were removed. However, the fITS dataset had a low average read number per sample, so samples below 1,000 reads were removed.

For the 16S dataset, ASVs with the same taxonomic assignment at genus level were agglomerated using the tax_glom function in the phyloseq package. For the pITS and fITS datasets, ASVs with the same taxonomic assignment at species level were agglomerated. Taxonomic agglomeration reduces the risk of artificially splitting species into separate ASVs and so overestimating alpha diversity (Schloss 2021). Taxonomic agglomeration reduced the species number drastically in the 18S and COI datasets, and so was not carried out.

The confidence of taxonomic assignment in the five datasets was variable. For the 16S and fITS dataset, all taxa with a root confidence assignment of <100 were filtered out. For the pITS and COI datasets, all taxa with an overall confidence of assignment of <0.97 were filtered out. The 18S dataset had a very low proportion of taxa with high confidence of assignment, so any taxa with root confidence <0.65 were filtered out. The results for these taxa should be interpreted in the light of this.

Finally, non-target taxa were filtered from the 16S dataset (assigned as “Mitochondria” and “Chloroplast”) and the pITS dataset (“Fungi” and “Unassignable”). Any field, extraction or PCR blanks that had not been removed from preceding steps were then removed.

The 16S blank samples contained 47 unique taxa, of which 43 were present in the pond water and sediment samples. One taxon accounted for 82 % of all reads in the blank samples (unclassified_Enterobacteriaceae), indicating possible laboratory contamination. However, this taxon did not appear as one of the most abundant in the environmental samples (see Figure S2.3). Two other taxa (unclassified Comamonadaceae) accounted for 5 %, but again were not abundant in environmental samples. The other taxa were present in total <150 reads across the four blanks indicating a low level of cross-contamination between environmental samples and blanks.

The 18S blank samples contained 11 unique taxa, and 10 of these were present in the pond water and sediment samples. None bar three were identified to genus level, and these had total reads in blanks <50. The most read-abundant taxa were unidentified below Domain level.

The fITS blank samples contained 5 unique taxa, and 3 of these were present in the pond water and sediment samples. One taxa, identified as genus *Tylospora* was present in environmental samples at high read abundances, whereas the other 2 taxa accounted for <200 reads across all samples. *Tylospora* reads were removed from environmental samples and subsequent analysis.

The COI blank samples contained 32 unique taxa, all of which were present in the pond water and sediment samples. Only 3 taxa were present in environmental samples above 200 reads across all samples, indicating cross-contamination in the laboratory stages. These were identified as *Homo sapiens, Hydropsyche siltalai* (a species of caddisfly) and *Nais elinguis* (aquatic annelid). These were not removed from the dataset due to likelihood of these species' eDNA being present in pond environments. Finally, the pITS blank samples contained 0 unique taxa.

Statistical analyses

All statistical analyses were carried out in R, version 4.2.0 (R Core Team 2021). To assess differences in abiotic variables between pond types, ANOVAs and Kruskal-Wallis tests were carried out using functions in base R. Abiotic variable plots were made by using ggplot2 (Wickham 2009).

Almost all other statistical analyses were carried out in the R package microeco, version 0.11.0 (Liu et al. 2021), for each of the five datasets separately. Relative abundance of different taxa was calculated, and bar graphs plotted, using functions within the trans_abund class. Taxa abundance between different sample types (sediment vs water) and pond types was compared via t.test and Kruskal-Wallis test using functions within the trans_diff class. Three measures of alpha diversity (Simpson, Shannon and Chao) were calculated and compared via t.test or Wilcoxon test within the trans_alpha class. Functional diversity was assigned to prokaryotic taxa in the 16S dataset and fungal taxa in the fITS dataset using the inbuilt function datasets in microeco. For 16S, this was FAPROTAX v1.2.4 (Louca et al. 2016), and for fITS this was FungalTraits ver 1.2 (Polme et al. 2020). Percentages of taxa of different functional groups were calculated using functions within the trans_func class.

To examine beta diversity, functions within the trans_beta class were used. Beta diversity was calculated using a Bray-Curtis index and a Principle Co-ordinates Analysis (PCoA) was performed using the cal_ordination function. A permutational multivariate analysis of variance (PERMANOVA) was performed using the cal_manova function, which uses the adonis2 function in the vegan package, version 2.6-4 (Okansen et al. 2022). This was performed on all samples to examine the effect of sample type, and pond type on community composition, including sample type, pond type and their interaction as factors in the analysis.

The same analysis of beta diversity and community composition (calculation of Bray-Curtis index, PCoA and PERMANOVA) was repeated on all five datasets after they had been

transformed from read-abundance to presence-absence data (all reads >0 replaced with 1). The five datasets were also remade so there was neither filtering taxa by confidence of taxonomic assignment, nor taxonomic agglomeration, and the analysis of beta diversity and community composition was repeated again.

To examine whether, for larger ponds, it is better to collect several water samples and merge in the field or sequence each sample separately, the Main Pond water samples were separated from the other samples for each of the five taxonomic communities. For the sample which was a mixture of water from 17 separate locations ("MPm"), the relative read abundance of different taxa was extracted, either at phylum or class level. This was compared to the mean read abundance of different taxa from the separate water samples (MP1-17).

To compare eDNA metabarcoding species detection with traditional visual-based monitoring methods, the final species list from the COI and pITS datasets were compared with macro-invertebrate and macrophyte records from the most recent survey of Pinkhill Meadow at the time of analysis (Williams 2017). All higher plant (Embryophyta) taxa in the pITS dataset (n=46) were assigned a 'wetland status' function according to the species list used in the National Pond Survey (Biggs et al. 1998). Floating-leaved and submerged species were combined into an "aquatic" category, whereas emergent species and trees or shrubs were combined into a "wetland" category. Plants which did not appear on the NPS list were classified as "terrestrial".

To compare how eDNA bioinformatic processing decisions affect the final beta diversity metrics, three different datasets were constructed for each ecological community (bacteria, eukaryotes, fungi, animals, and green plants) using different approaches to handle sequence read filtering. One dataset did not filter out taxa which had a low confidence of taxonomic assignment, but each taxa in this dataset was still weighted by read abundance (the "unfiltered dataset"). The second dataset was the standard dataset, filtered by confidence of taxonomic assignment and read-abundance weighted (the "filtered dataset"). The third dataset was filtered, and read abundance weighting was removed (the "presence-absence dataset"). The unfiltered 16S dataset contained 791 taxa, the unfiltered 18S dataset contained 2,474 taxa, the unfiltered fITS dataset 2,469 taxa, the unfiltered COI dataset 5574 taxa and the unfiltered pITS dataset 671 taxa.

Results

eDNA concentrations, ASV and read numbers

After filtration and DNA extraction, the concentration of DNA from pond water samples ranged from –1.3 µg/ml to 35.1 µg/µl (*M* = 9.84, *SD* = 10.35). In sediment samples, the concentration after extraction ranged from 1.2 µg/ml to 105.4 µg/µl (*M* = 22.31, *SD* = 29.27).

After processing with the DADA2 pipeline (quality trimming, de-replication, chimera removal), the raw datasets were as follows: for 16S sequences, 7,597 ASVs, and the total number of reads across 72 samples was 1,492,081. For 18S sequences, 17,831 ASVs and 1,922,651 reads across 79 samples. For COI, the total number of ASVs was 18,480 and there were 1,418,542 reads across 80 samples. For fITS, the total number of ASVs was 7,613 and there were 1,587,404 reads across 70 samples. Finally, for pITS, this process outputted 1,536,985 raw reads in 2,876 ASVs across 71 samples.

Pond abiotic conditions

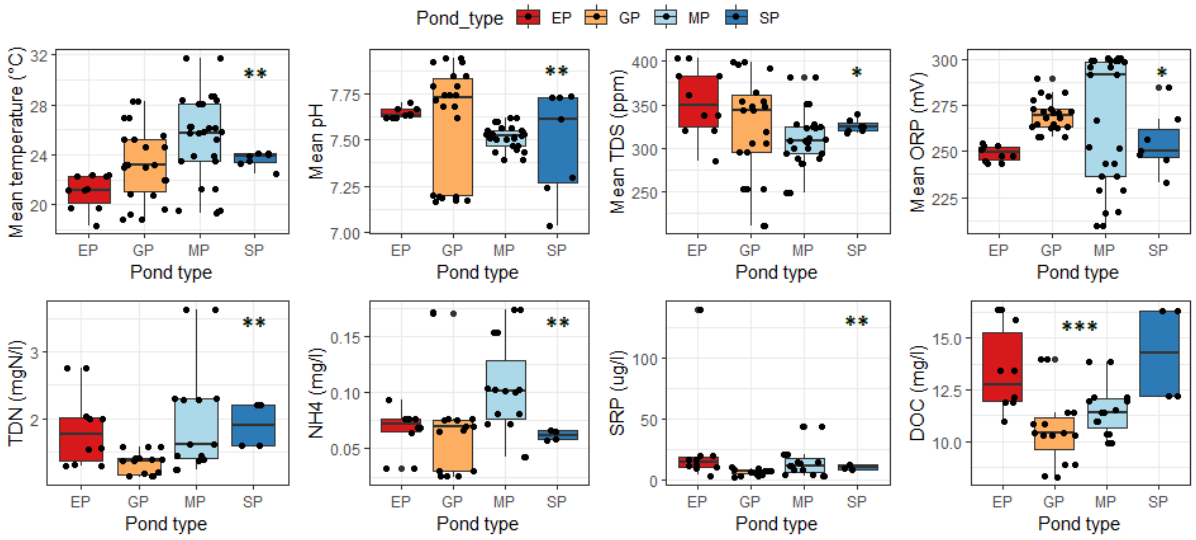


Figure 2.2: Kruskal-Wallis one way analysis of variance test results as a boxplot for each physiochemical variable by pond type. One, two and three asterisks indicate that the analysis was significant at the p<0.05, p<0.01 and p<0.001 levels respectively.

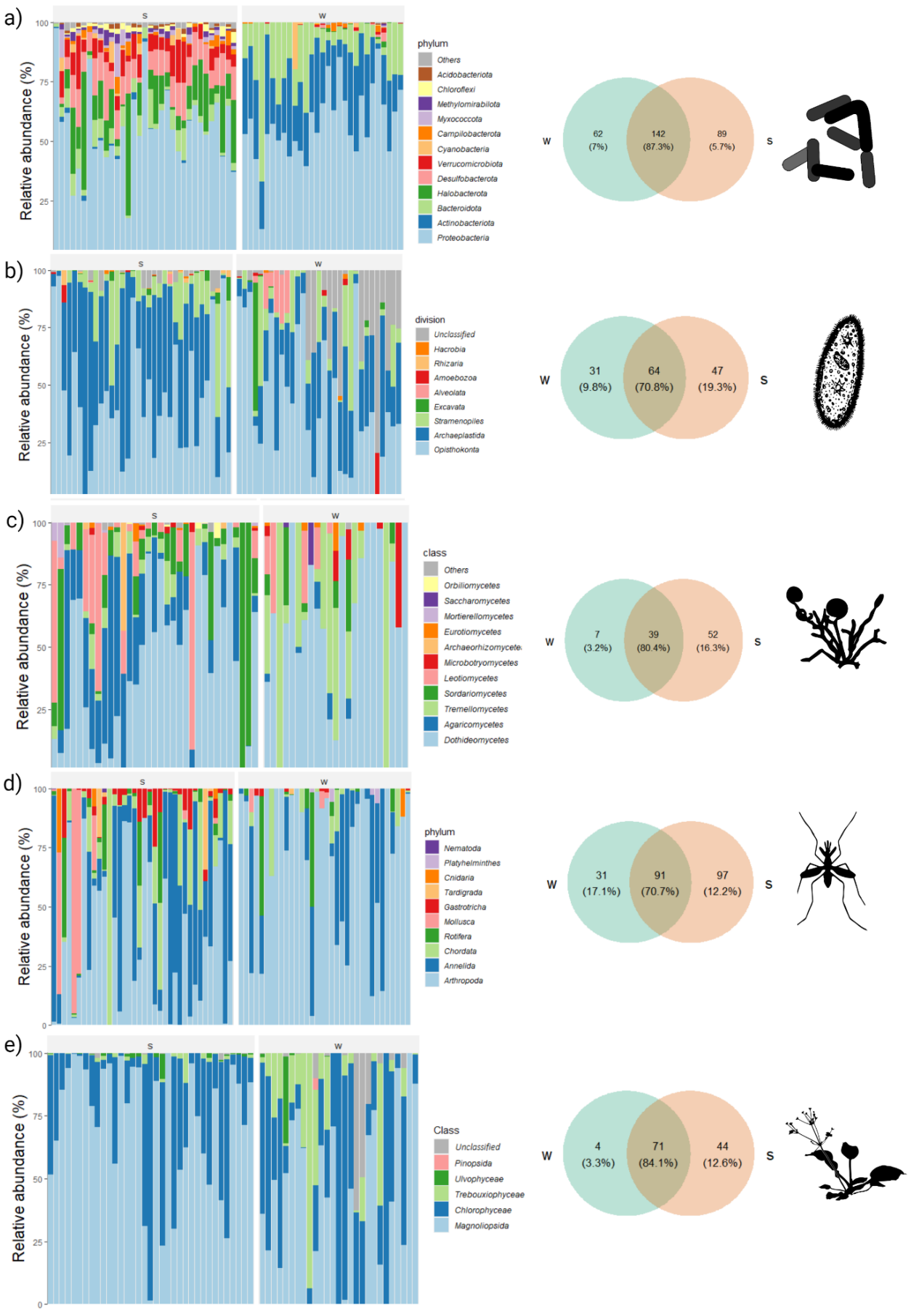
The water and sediment samples were collected on 2nd June 2020, at the end of a spring which had the most sunshine hours since records began, and was one of the driest on record (Met Office 2020). The River Thames level was 0.853 m (River Levels UK 2023),

typical for the time of year. The site was not flooded, and some sites that had been sampled in previous sampling events (see Chapter 3) were dry (MP1, MP2, MP4 and EP4).

Kruskal-Wallis tests were performed to analyse the differences in physiochemical variables between the different pond types. No significant differences were found for pond depth (H (3) = 4.32, p = 0.23) or canopy cover (H(3) = 6.97, p = 0.073). Significant differences between pond types were found for eight other measured variables (pH, H(3) = 12.75, p = 0.005, temperature, H(3) = 15.07, p = 0.002, TDS, H(3) = 8.23, p = 0.042, ORP, H(3) = 10.43, p = 0.015, TDN, H(3) = 13.18, p = 0.004, NH4, H(3) = 11.75, p = 0.008, SRP, H(3) = 13.40, p = 0.004, DOC, H(3) = 17.47, p < 0.001, see Figure 2.2). Results of post-hoc Dunn's tests with a Bonferroni correction are reported in Table S2.2.

Taxonomic and functional diversity in Pinkhill Meadow ponds

Figure 2.3 (below): Taxonomic composition and number of taxa in sediment and water samples for a) Bacteria and Archaea (16S dataset), b) Eukaryota (18S dataset), c) Fungi (fungal ITS2 dataset), d) Animalia (COI dataset), and e) green plants and algae (plant ITS2 dataset). In bar graphs, "s" = sediment samples, "w"= water samples, and the y axis shows relative eDNA read abundance as a percentage. In Venn diagrams, the number of taxa is above and the % proportion those taxa contributed to overall read abundance below in brackets.


a)
S
W
Relative abundance (%)
100
75
50
25
phylum
Others
Acidobacteriota
Chloroflexi
Methylomirabilota
Myxococcota
Campilobacterota
Cyanobacteria
Verrucomicrobiota
Desulfobacterota
Halobacterota
Bacteroidota
Actinobacteriota
Proteobacteria
W
62
(7%)
142
(87.3%)
89
(5.7%)
S
b)
S
W
Relative abundance (%)
100
75
50
25
division
Unclassified
Hacrobia
Rhizaria
Amoebozoa
Alveolata
Excavata
Stramenopiles
Archaeplastida
Opisthokonta
W
31
(9.8%)
64
(70.8%)
47
(19.3%)
S
c)
S
W
Relative abundance (%)
100
75
50
25
class
Others
Orbiliomycetes
Saccharomycetes
Mortierellomycetes
Eurotiomycetes
Archaeorhizomycetes
Microbotryomycetes
Leotiomycetes
Sordariomycetes
Tremellomycetes
Agaricomycetes
Dothideomycetes
W
7
(3.2%)
39
(80.4%)
52
(16.3%)
S
d)
S
W
Relative abundance (%)
100
75
50
25
0
phylum
Nematoda
Platyhelminthes
Cnidaria
Tardigrada
Gastrotricha
Mollusca
Rotifera
Chordata
Annelida
Arthropoda
W
31
(17.1%)
91
(70.7%)
97
(12.2%)
S
e)
S
W
Relative abundance (%)
100
75
50
25
0
Class
Unclassified
Pinopsida
Ulvophyceae
Trebouxiophyceae
Chlorophyceae
Magnoliopsida
W
4
(3.3%)
71
(84.1%)
44
(12.6%)
S

a. Bacteria and Archaea (16S rRNA gene)

After bioinformatic processing, the final 16S dataset consisted of 293 taxa across 62 samples (sediment n = 33, water n = 29). Nearly half of all taxa (142) were found in both water and sediment samples, however, these 142 taxa made up around 87 % of total read abundance (see Figure 2.3a). Proteobacteria was the most abundant phylum, with a mean of 54.5 % (*SD* = 15.4 %) of water samples' and 57.4 % (*SD* = 17.2 %) of sediment samples' read abundance. Bacteroidota was also common, comprising a mean 16.0 % (*SD* = 13.7 %) of water samples and mean 5.4 % (*SD* = 3.8 %) of sediment samples. Actinobacteria was prominent in water samples, with a mean of 24.2 % (*SD* = 10.5 %) of total read abundance, but was virtually absent from sediment samples (*M* = 0.6 %, *SD* = 0.8 %). Most other phyla were more abundant in sediment samples than water samples, including Desulfobacterota, Verrucomicrobiota, Methylmirabilota, Chloroflexi, Acidobacteriota and the Archaen phylum Halobacterota (see Figure 2.3a).

The top 25 bacterial and archean genera by read abundance made up around 50 % of the total reads in water samples, and 40 % in sediment samples (see Figure S2.3). The most abundant genera in sediment samples were *Methanoregula, Methanosaeta, Luteolibacter, Candidatus accumilibacter, Sva0081 sediment group, Geobacter* and *Crenothrix.* This small selection highlights the large functional diversity of freshwater pond sediments: *Methanoregula* and *Methanosaeta* are both methanogens (Brauer et al. 2006, Patel & Sprott 1990), whereas *Crenothrix* is a methane-oxidising genus (Oswald et al. 2017). *Candidatus accumilibacter* is a key organism in phosphate-removing wastewater treatment (Seviour et al. 2003), *Geobacter* species are known to reduce iron and manganese oxides and aromatic hydrocarbons (Lovley et al. 2011), and the *Sva0081 sediment group* are all sulphate-reducing bacteria (Fonseca et al. 2022).

In water samples, the most abundant genera were *Pseudarcicella, Flavobacterium, Polynucleobacter, hgcI clade, Planktoluna, Rickettsiella* and *Novosphingobium.* The functions of these genera are not always well described, but they include potential endosymbiotic organisms; *Polynucleobacter,* in ciliates (Heckman & Schmidt 1987), and *Rickettsiella* in arthropods (Jurat-Fuentes & Jackson 2012); and *Pseudarcicella* was isolated from the skin of medicinal leeches (Kampfer 2012). *Flavobacterium, Polynucleobacter* and the *hgcI clade* have previously been associated with urban-impacted waters and nitrogen-rich environments (Adyasari et al. 2020), whereas *Novosphingobium* has been associated with the biodegradation of aromatic compounds (Gan et al. 2013).

According to the functional assignment using FAPROTAX (see Figure S2.1), the majority of taxa in all samples were aerobic chemoheterotrophs. Compared with water samples, sediment samples had a greater proportion of taxa which were anaerobic chemoheterotrophs, or were responsible for carbon cycling as fermenters, methanogens, methanotrophs or methylotrophs. Taxa involved in sulphur cycling were more prevalent in sediments than the water column. Taxa with nitrogen, iron and manganese cycling capabilities were found in several samples of both type.

b. Eukaryotes (18S rRNA gene)

The final 18S dataset consisted of 142 taxa across 65 samples (sediment n = 34, water n = 31). Around 45 % of taxa were found in both types of sample, and these made up around 70 % of total read abundance. Sediment samples had more unique taxa and a greater proportion of overall read abundance than water samples (see Figure 2.3b). Most taxa detected were either Opishtokonta (45) or Archaeplastida (41), with the remainder Stramenopiles (17), Excavata (12), Amoeobozoa (6), Rhizaria (3), Hacrobia (3), or Alveolata (1). Both types of sample had a similar proportion of Opisthokonta reads (sediment, *M* = 43.9 %, *SD* = 26.5 %, water, *M* = 43.3 %, *SD* = 28.7 %). Sediment samples contained significantly more Archaeplastida reads than water samples (sediment, *M* = 40.8 %, *SD* = 25.1 %, water, *M* = 25.6 %, *SD* = 17.6 %, see Figure 2.3b). Water samples also contained a significantly higher proportion of reads from taxa unclassified to Division level (*M* = 15.7 %, *SD* = 20.6 % compared to *M* = 2.0 %, *SD* = 3.1 % in sediment samples).

Only 32 taxa were identified to genus level in the 18S dataset, and only 84 to class level. The most read-abundant classes in sediment samples were Gastrotricha, Annelida, Mollusca, Charophyceae (“stoneworts”, large algae which are traditional indicators of clean freshwater and surveyed using visual techniques) and Bacillariophyta (diatoms). In water samples, the most read-abundant class was Arthropoda, making up between 15 % - 40 % of all reads, followed by Gastrotricha, Trebouxiophyceae, and Euglenida (see Figure S2.4).

c. Fungi (fungal ITS2)

The final fITS dataset contained 98 taxa across only 56 samples (sediment n = 33, water n = 23). Around half of all taxa (52) were found uniquely in sediment samples, comprising 16.3 % of all reads, whereas the 39 taxa found in both sediment samples and water samples were more common, making up 80.4 % of all reads (see Figure 2.3c). The most common class in both sediment and water samples were Dothideomycetes (water, *M* = 57.0 %, *SD* = 32.1 %, sediment, *M* = 43.4 %, *SD* = 32.1 %). Agaricomycetes were significantly more

abundant in sediment samples than water samples (sediment, *M* = 21.9 %, *SD* = 22.4 %, water, *M* = 1.5 %, *SD* = 3.4 % Tremellomycetes were significantly more abundant in water samples (*M* = 23.1 %, *SD* = 22.1 %) than sediment samples (*M* = 3.7 %, *SD* = 4.5 %). Sodariomycetes and Leotiomycetes both made up around 12 % of sediment sample reads (Sodariomycetes: *M* = 11.9 %, *SD* = 18.0 %, Leotiomycetes: *M* = 13.9 %, *SD* = 21.1 %), and less than 3 % of water sample reads (Sodariomycetes: *M* = 2.9 %, *SD* 7.0 %, Leotiomycetes: *M* = 1.9 %, *SD* 3.8 %).

The top 25 most abundant fungal genera composed around 75 % of total read abundance in sediment samples, and slightly less than this, 65 %-75 % in water samples (see Figure S2.5). The most abundant genera in sediment samples were*, Stagonospora*, many species of which are major plant pathogens (JGI Mycocosm 2023), *Tomentella, Pseudeurotium,* and *Helicodendron,* an aero-aquatic hyphomycete (Glen-Bott 1951). The most abundant genera in water samples included *Vishniacozyma*, *Alternaria, Cladosporium* and *Beauvaria* *Alternaria* are common and widespread plant pathogens and can cause infections in immunocompromised people (Patriarca et al. 2014), *Cladosporium* are common indoor and outdoor moulds (Stewart & Robinson 2013) and many *Beauvaria* species are insect pathogens (Wang et al. 2022). Sediment samples had a significantly higher alpha diversity than water samples for all three metrics tested (Shannon, Chao and Simpson diversity, t. tests, all $p<0.01$).

Initial functional diversity analysis of the 99 taxa (see Figure S2.2) indicated a broad range of life forms (e.g., unicellular yeasts, multicellular mychorrizal species) and lifestyles (saprotrophs, plant pathogens, animal parasites, symbiotic ectomychorrizal taxa, foliar endophytes). In total, only 6 taxa were classed as “aquatic”, 41 taxa “partly aquatic” and 25 taxa “non-aquatic”. However, the majority of reads in each sample were from those fungal taxa classed as aquatic. There appeared to be no clear differences in functional diversity between sediment and water samples.

d. Metazoans (COI mtDNA)

The COI dataset contained 234 taxa across 70 samples (n = 36 sediment, n = 24 water). Again, sediment samples contained more unique taxa (87) than in water samples (35), but the taxa shared between the two sample types made up the majority of total read abundance (104, 59 %). Water samples had a much higher proportion of Arthropoda reads than sediment samples (water, *M* = 52.9 %, *SD* = 34.4 % compared to sediment, *M* = 19.3 %, *SD* = 24.2 %), whereas sediment samples had a higher proportion of Annelida reads (sediment, *M* = 32.8 % *SD* = 31.8 %, water, *M* = 12.0 %, *SD* = 24.3 %). Sediment samples also

contained more Bacillariophyta (diatom) and Gastrotricha reads, a mean of 9.1 % (*SD* = 12.5 %), and 4.7 % (SD 6.4 %) respectively. Water samples contained more reads assigned to Cryptophyta, (*M* = 10.9 %, *SD* = 18.6 %). Dinoflagellata (called Pyrrophycophyta in the graph) made up a substantial proportion of overall reads: mean 13.0 % (SD 16.3 %) in water samples, and 16.7 % (SD 31.5 %) in sediment (see Figure 2.3d).

The 25 most read abundant species in sediment and water samples made up between 60 and 95 % of the total read abundance (see Figure S2.6). In sediment samples, oligochaete annelid worms in the Naididae , such as *Nais communis, Dero obtusa, Dero digitata* and *Chaetogaster,* were common, making up 20 – 45 % of total reads. Two oomycete plant pathogens were common, *Phytopythium* and *Phytopthora.* The pond snail *Radix balthica* was the third most common species encountered. The eDNA of the most abundant species was assigned as *Alexandrium minutum* - likely a misassignment as this is a dinoflagellate species usually considered marine, and responsible for “red tides”.

In water samples, eDNA of small crustaceans was common, such as Copepod species/genera *Mesocyclops leuckarti, Thermocyclops* and *Eucyclops,* and Ostracods *Cypridopsis vidua,* and *Limnocythere.* Insects included a springtail species often found on the surface of ponds and puddles *Sminthurides aquaticus. Lingulodinium,* a dinoflagellate genus, was also present, of which the only extant species is bioluminescant. Additionally, the eDNA of two terrestrial species were found in water samples: *Lampyris noctiluca,* the glow worm (which, despite the name, is a beetle), and *Agapeta hamana* a species of moth whose food plant is thistle.

e. Plants and green algae (plant ITS2)

The final pITS dataset contained 119 taxa across 62 samples (n sediment = 35, n water = 27). Sediment samples again contained a higher number of unique taxa and a larger proportion of total read abundance than water samples. At the phylum level, 72 taxa were Chlorophyta (green algae) and 47 were Streptophyta (“higher” plants; Embryophytes). Sediment and water samples both had a similar mean proportion of reads assigned to the green algal class Chlorophyceae (*M* = 35.4 %, *SD* = 29.9 % in sediment and *M* = 26.1 %, *SD* = 25.5 % in water). However, sediment samples contained a far larger proportion of Magnoliopsida (dicotyledons or flowering plants) than water samples: a mean of 71.9 % (SD

26.6 %) of reads compared to mean 42.1 % (SD 32.3 %) of reads. Water samples contained a greater proportion of algal class Trebuxiophyceae and of sequences unclassified at the class level (see Figure 2.3e).

The 25 most abundant plant species made up between 40 % and 75 % of overall read abundance in samples (see Figure S2.7). The most read abundant species in sediment samples included *Chlamydomonas petasus, a* unicellular chlorophyte alga, *Phragmites australis* (southern water-reed), *Potamogeton berchtoldii* (small pondweed), *Alisma plantago-aquatica* (water plantain), *Sparganium stoloniferum* (a bur-reed species) and *Carex acutiformis* (lesser pond sedge). In water samples, the most read abundant species were primarily chlorophyte algae such as *Chlorococcum* and *Chlamydomonas* species, and the cultivated species *Brassica oleracea* and *Latuca serriola.*

Of the 47 higher plant (Embryophyta) species, 13 were classed as “wetland”, 30 “terrestrial” and only 4 “aquatic”, according to the National Pond Survey species list (see methods). Additionally, 12 species were those which are cultivated (e.g., *Arachis hypogaea,* Peanut) or have wild and cultivated varieties (*Brassica* sp.). In water samples, 70.3 % of the total eDNA reads identified as from terrestrial species, with 22.7 % coming from wetland species and 7.0 % from aquatic species. By contrast, 66.2 % of the eDNA reads in sediment samples came from wetland species, 15.8 % from aquatic species and only 18.0 % from terrestrial species (see Figure 2.4). In water samples, nearly a third (32.5 %) of eDNA came from cultivated species, whereas in sediment samples only 5.83 % of eDNA did.

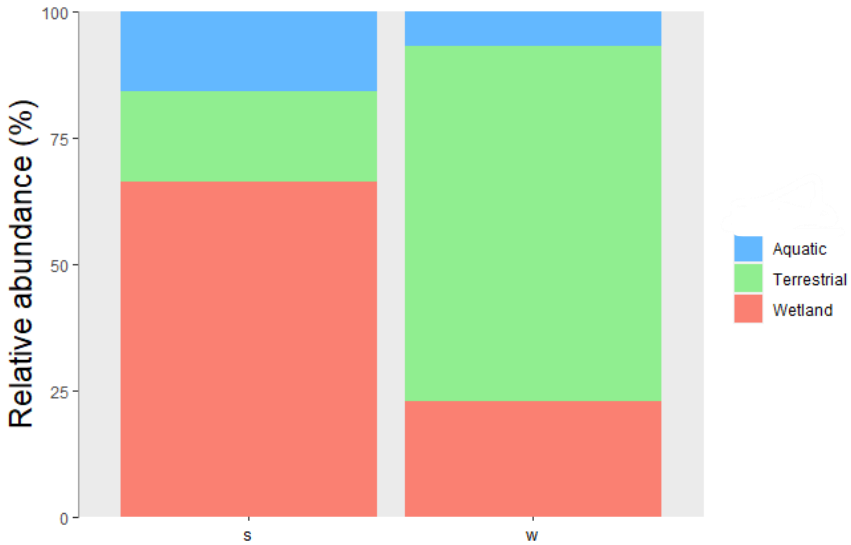


Figure 2.4: Proportion of total eDNA read abundance originating from wetland, terrestrial and aquatic higher plant species. "s"= sediment samples (n=35), "w" = water samples (n=27). Total number of species = 47.

Community composition: unfiltered, presence-absence and read abundance-weighted data.

For all PERMANOVA analyses of the five different taxonomic communities, sample type (sediment or water) and pond type (main pond, surface-water pond, experimental pond or groundwater pond) were always significant factors influencing community composition, no matter whether unfiltered, filtered, or presence-absence, datasets were used (see Table 2).

The proportion of overall variance explained by sample type was much higher for bacteria than for the other four groups (fungi, eukaryotes, animals and green plants and algae). In bacterial communities, sample type accounted for 30 % - 35 % of the overall variance in community composition, whereas it accounted for between 6.3 % and 14 % of the variance in community composition of the four other communities (see Table 2.2). Furthermore, for prokaryotic organisms, the sample type was a much more influential factor than the pond type, explaining over five times the overall amount of community variation, whereas for the other groups, sample type and pond type appeared to have a roughly equal effect on community composition.

The PERMANOVAs using presence-absence datasets showed more pronounced patterns than the read-abundance weighted datasets, with the R2 values of most factors in the presence-absence analyses being greater than the R2 values of the same factors in the read-abundance weighted analyses (see Table 2.2). The exception to this was the sample type - pond type interaction factor, which in some cases was significant in the read abundance-weighted datasets and not significant in the presence-absence datasets.

Compared to the read abundance-weighted datasets, the analyses using unfiltered datasets did not show a clear pattern: for instance, for bacteria, sample type had a smaller effect ($R2$ = 0.30 unfiltered, $R2$ = 0.34 filtered) but for animals, it had a greater effect ($R2$ = 0.072 unfiltered, $R2$ = 0.035 filtered). The interaction between sample type and pond type was always significant in the unfiltered datasets, whereas this was not true of other datasets.

Table 2.2: R2 values of factors in PERMANOVA analyses across five taxonomic groups in Pinkhill Meadow ponds, June 2020. The first three columns are results from datasets without filtering by confidence of taxonomic assignment. The middle three columns are read abundance-weighted datasets, and the latter three results from presence-absence (0 or 1) datasets. In all cases, the degrees of freedom were: "Type", 1, "Pond type", 3 and "Interaction", 3. One, two and three asterisks indicate whether the analysis was significant at the p<0.05, p<0.01 and p<0.001 level respectively.

	Unfiltered			**Filtered**			**Presence/absence**		
	Type	*Pond type*	*Interaction*	*Type*	*Pond type*	*Interaction*	*Type*	*Pond type*	*Interaction*
16S Bacteria & Archaea	0.30 ***	0.062 **	0.063 **	0.34 ***	0.058 *	0.059 *	0.35 ***	0.053 *	0.042
18S Eukaryotes	0.063 ***	0.089 ***	0.078 ***	0.065 ***	0.091 ***	0.068 ***	0.118 ***	0.130 ***	0.073 ***
COI Animalia	0.072 ***	0.083 ***	0.061 ***	0.035 ***	0.059 ***	0.046	0.076 ***	0.092 ***	0.041
fITS Fungi	0.072 ***	0.099 ***	0.047 **	0.089 ***	0.078 **	0.035	0.137 ***	0.088 ***	0.037
plTS Green plants & algae	0.058 ***	0.090 ***	0.046 ***	0.086 ***	0.084 ***	0.045 *	0.137 ***	0.099 ***	0.036

Merged and separate water samples and community representation

In order to determine whether spatial heterogeneity of eDNA could influence biodiversity estimates from larger ponds, 16 separate water samples were collected from around the perimeter of the main pond and processed separately, alongside a single merged sample combining water from the 16 separate locations. The 17 samples all had the same volume of water filtered (see Methods). The samples went through the standard bioinformatic pipeline (filtered and read abundance-weighted) as outlined in the Methods. After extracting the Main Pond samples from the full datasets, the samples from the 18S dataset and the fITS dataset were discarded due to low taxa number (62 for the 18S dataset and 20 for the fITS dataset) and low read abundance. This left the 16S Bacteria dataset, the COI Animalia dataset and the ITS2 green plants and algae dataset, for which the relative read abundances of different

taxa were compared between the separate samples and the merged sample. The results are plotted in Figure 2.5.

For the bacteria dataset, the merged sample and the mean of the read abundance in separate samples broadly matched up, especially for the most abundant phyla (Gammaproteobacteria: 62.49 % of merged total reads, *M* = 64.88 %, *SD* = 14.53 % of separate reads; Actinobacteria: 20.85 % merged, *M* = 23.57 %, *SD* = 12.09 % separate, see Figure 2.5 and table S2.3). However, for both the Animalia dataset and the green plants and algae dataset, the mean proportion of classes from the separate samples did not match the proportion of different classes found in the merged sample (see Figure 2.5). For instance, Ostracoda made up only 2.5 % of reads in the merged sample, but composed a mean of 15.1 % (SD 23.8 %) of reads in separate samples. Conversely, Apiales reads appeared at a high percentage in the merged water sample (41.72 %) but were not found in any of the separate samples.

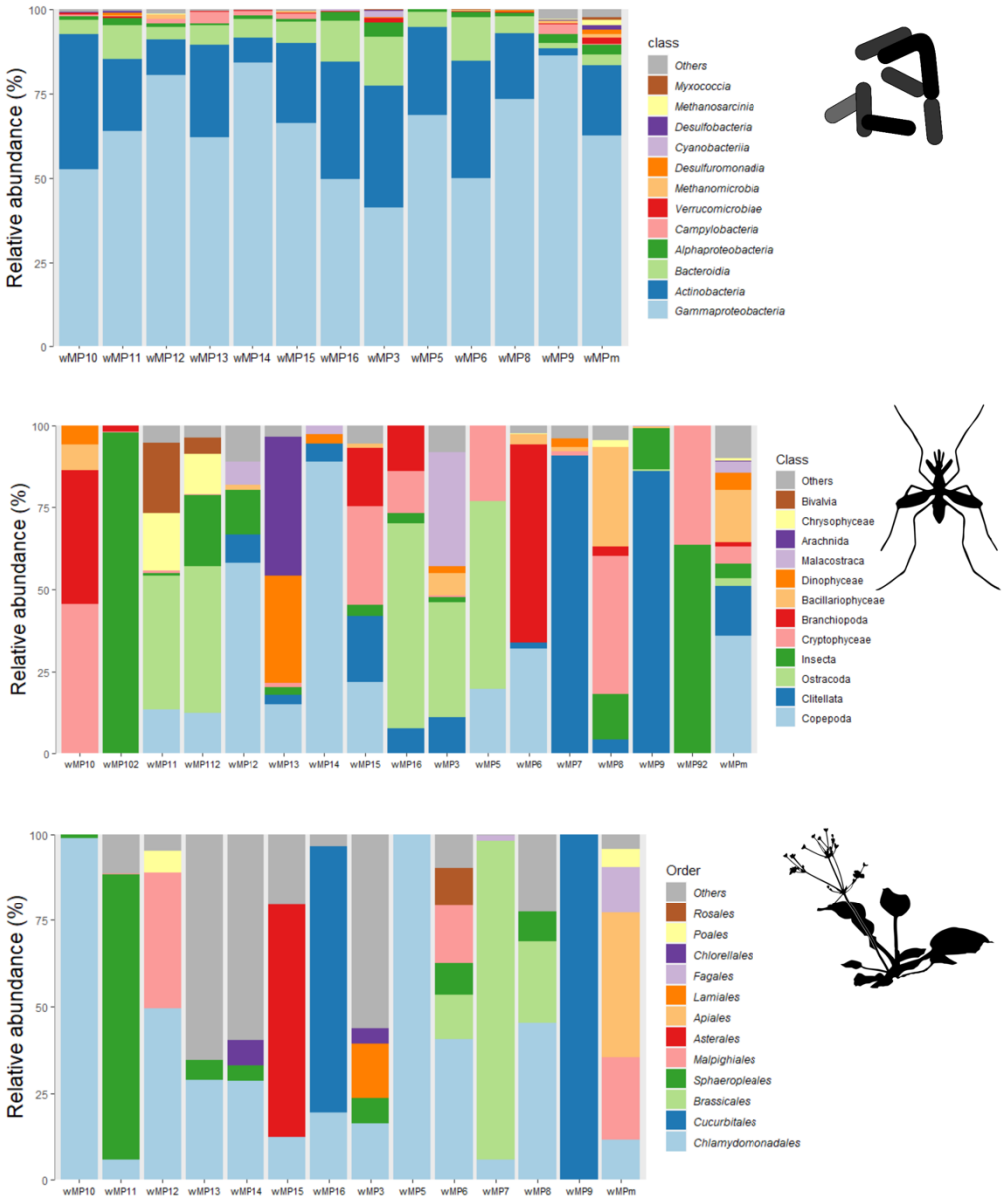


Figure 2.5: Taxonomic composition of water samples from the Main Pond of Pinkhill Meadow for a) Bacteria b) Animalia and c) Green plants and algae. In each plot, the bar furthest to the right (labelled "MPm") is the water sample merged in the laboratory, the other bars are the separate water samples.

Overlap between traditional survey methods and eDNA survey methods

At the most recent full plant survey of the Pinkhill Meadow wetland complex carried out in 2016 (Williams 2017) recorded 78 aquatic or emergent higher plant species (Embryophytes)

across the entire site, noting that this is “exceptionally rich in plant species, supporting approximately 20 % of all Britain’s freshwater taxa”. By comparison, my eDNA survey carried out in June 2020 detected 47 Embryophyta species, and only 8 species were detected with both the eDNA and the visual survey method (see Figure 2.6).

Species detected using eDNA but not detected visually were mainly either terrestrial species or trees which were not recorded by Williams in 2016. However, some wetland species were detected, such as *Sparganium stoloniferum* (a bur-reed native to East Asia), *Stuckenia pectinata* (ribbon weed) and *Alopecurus aequalis* (orange foxtail). These may have been misidentified (as in the case of *Sparganium stoloniferum,* which probably originated from one of the two other *Sparganium* species on-site) or perhaps simply missed visually.

Some taxa, such as a *Carex, Salix, Scheonoplectus* and *Eleocharis* species were not identified to species level. When looking at genus level, visual methods identified 59 different higher plant genera and eDNA methods 29, with an overlap of 13 genera.

There were only three overlaps at the species level between Animalia taxa detected using visual-based survey methods and the eDNA metabarcoding method using the COI primer pair: R*adix balthica* (pond snail), *Hydroporus planus,* a predaceous diving beetle, and *Noterus clavicornus,* a burrowing water beetle.

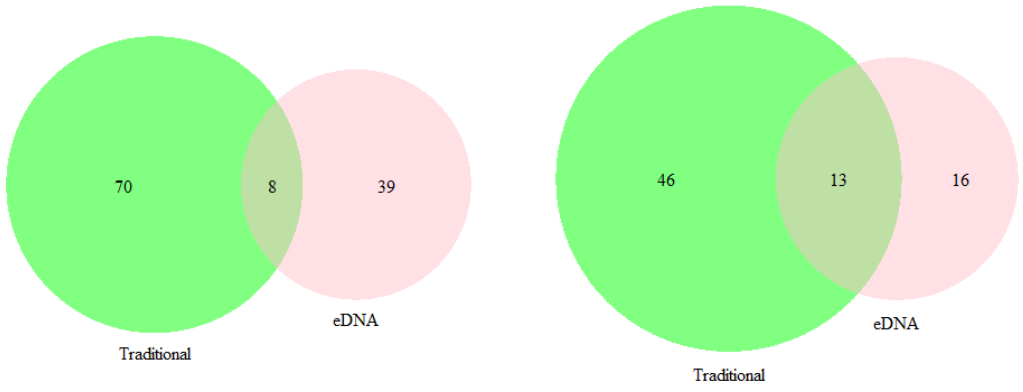


Figure 2.6: Overlap between left, aquatic and wetland macrophyte species and right, genera detected using traditional, visual-based survey methods and eDNA metabarcoding in Pinkhill Meadow ponds, June 2020.

Discussion

This study has shown that eDNA metabarcoding using multiple taxonomic markers is a tool with great potential to shed new light on the ecology of freshwater ponds and to monitor their health, but careful selection of the methods used in the field, the laboratory and *in silico* is required.

Use of multi-taxa metabarcoding for surveying pond ecosystems

The sequencing of amplicons from five different primer pairs on a single v3 flow cell using the Illumina MiSeq platform had not, to our knowledge, been carried out previously. The process was a success, with all five different amplicons producing high total read numbers and broadly consistent reads across samples. Our results reveal that an extraordinary level of biodiversity was found in these 20 ponds on a single sunny June day: the final (confidence-filtered and taxonomically agglomerated) dataset contained 887 taxa, nearly five times the number of species detected using traditional survey methods in 2016 (181: 103 macroinvertebrate species and 78 macrophyte species, Williams 2017).

All five of the assays predominantly detected taxa which are not monitored traditionally, covering a much broader swathe of the tree of life than is usually surveyed in ponds (or in many other natural environments). These included meiobenthos such as Nematoda and Gastrotricha (COI, 18S), macrobenthos such as Annelida, Mollusca and Platyhelminthes (COI, 18S), zooplankton such as Ostracoda, Copepoda and Malacostraca (COI), many different algal groups such as Dinoflagellates, Bacillariophyta, Chlorophyta and Cryptophyta (18S, COI, pITS), protozoans such as Euglenida (18S), and numerous different fungal (fITS) and prokaryote (16S) classes. All these taxonomic groups are difficult and time-consuming, if not impossible, to identify using microscopy, and therefore are not systematically monitored in pond ecosystems that we know of. In addition, the assays detected ‘higher’ plants (Embryophytes), both aquatic and terrestrial (pITS), and some larger insect and Arachnida taxa (COI).

Initial analysis of functional prokaryote diversity revealed taxa involved in the cycling of nitrogen, sulphur, metals and carbon, including both methane respiration and methanogenesis (see Figure S2.1). This suggests future research into the ecosystem functions of freshwater ponds using eDNA of prokaryotes, as has been carried out on larger freshwater bodies (e.g., Zhang et al. 2018). For instance, there is current interest into whether, and under what conditions, small freshwater bodies become atmospheric carbon

sources or sinks (Rosentreter et al. 2021). There is also great potential to use prokaryotic eDNA to assess anthropogenic impact on freshwater ponds, as has been carried out in wetlands (e.g., Wood et al. 2021) and lakes (e.g., Liu et al. 2022).

Although these five assays combined detected a broad taxonomic extent, some of the assays struggled to identify taxa to lower taxonomic levels, for instance the 18S rRNA assay, where taxonomic assignment was low in many cases (66 % of ASVs in sediment samples and 88 % of ASVs in water samples unassigned at the Division level). Although subsequent assays using a different 18S rRNA primer pair in the same environment have had more success (see Chapter 3), this does point to a general feature of universal primers: that it is difficult to achieve both taxonomic breadth and specificity (Francioli et al. 2021). Incomplete reference databases reduced the amount of information in some cases: for instance, the confidence of taxonomic assignment was also low for many fungal ASVs, reducing the dataset from 7,613 ASVs to 99. This is to be expected as it is estimated that less than 7 % total fungal diversity has been described (Hawksworth & Lucking 2017), and aquatic fungi are particularly overlooked (Ittner et al. 2018).

The taxonomic composition detected by the COI primer pair and the 18S primer pair overlapped (both groups detected Arthropoda, Mollusca, Annelida, Bacillariophyta and Rotifera eDNA), but each primer pair also detected taxa the other did not (e.g., Nematoda for 18S, Dinoflagellata (Pyrrophycophyceae) for COI.) Other studies have found differences in the taxa detected using 18S and COI assays on the same samples (Giebner et al. 2020, Horton et al. 2017). Similarly, the 18S primer pair detected Charophyceae (stoneworts, a traditional indicator taxon of clean water ponds), but these were not detected in the pITS assay. These results show how primer bias can significantly influence biodiversity assessments using eDNA metabarcoding, and the power of using multiple taxonomic markers.

eDNA metabarcoding using the COI marker gene identified only three of the 103 macroinvertebrate species identified using sweep-netting at Pinkhill Meadow in 2016 (Williams 2017). This is comparable to other eDNA metabarcoding/traditional comparisons in ponds (Harper et al. 2021) and may be due to a variety of reasons: greater volumes of meiofaunal eDNA compared to macrofaunal eDNA, less shedding of eDNA by hard-bodied organisms, and primer bias (Leese et al. 2021).

eDNA metabarcoding using the plant ITS2 marker gene identified eight of the 78 macrophyte species identified visually at Pinkhill Meadow, often only detecting a single species of each genus (e.g., rushes *Juncus,* sedges *Carex* and pondweeds *Potamogeton*). This is similar to

the overlap found by Drummond et al. (2021) of 13 %, using ITS2 in lake water, but lower than the 31 % reported by Alsos et al. (2018), using the trnL marker on lake sediment samples. I suspect this result is at least partly due to incomplete reference databases for the target aquatic and wetland plant taxa: for instance, the taxon identified as *Sparganium stoloniferum,* a bur-reed native to East Asia, was most probably in reality eDNA from one of our native *Sparganium* species. In addition, different plant taxa may shed eDNA at different rates. The findings may also be due to the high proportion of eDNA from terrestrial species which were found in the samples, which could have swamped the weaker signal from the aquatic and wetland species. Most plant species identified were terrestrial species (30 out of 47) and reads from terrestrial plants made up 70 % of the total read abundance in water samples, and 18 % of sediment samples.

Terrestrial species were not only found in the pITS dataset: terrestrial insects were detected in pond water using the COI assay (Figure S2.6), and upwards of 25 % of fungal taxa sequenced were assigned as terrestrial or only partly aquatic (Figure S2.2). This shows the ability for eDNA sampled from pond water and pond sediments to originate from the surrounding landscape and air, and travel to the pond via wind, water or on animal vectors, as has been shown previously in rivers (Reji Chacko et al. 2023, Yang et al. 2021, Matsuoka et al. 2019) and lakes (Alsos et al. 2018, Bista et al. 2017). This land-air-water eDNA linkage may be particularly pronounced in ponds due to their small size, and (often) large proportion of the water surface overhung by trees and other vegetation, at least compared to lakes. This feature of ponds could be a disadvantage for eDNA metabarcoding for biomonitoring of the pond environment itself, but conversely, an opportunity to use ponds as passive eDNA samplers of the surrounding terrestrial landscape (Ushio et al. 2017, Harper et al. 2019c, Sales et al. 2019).

Finally, for all five assays the community composition in water and sediment samples was distinctly different (see Figure 2.3 and Table 2.2), with between 40 % and 60 % of ASVs only detected in one of the sample types. This reflects the different ecological communities that inhabit the two pond microhabitats, as well as the different transport, retention and degradation of eDNA in sediments vs water.

The effect of taxonomic filtering and read abundance weighting on community composition metrics derived from eDNA metabarcoding

Overall, it appears from my data that community composition metrics derived from eDNA metabarcoding data are relatively robust to taxonomic filtering by confidence and read-

abundance weighting. Similar overall patterns of community composition were found for microbes, micro eukaryotes, fungi, plants and animals, no matter whether low confidence assignments were filtered out, or whether a presence-absence matrix was used instead of eDNA read abundance. This is contrary to previous findings (e.g., Machler et al. 2021), and points to an increased confidence in filtering methods, and using eDNA read abundance in samples as a proxy for taxon abundance in the environment.

In each case, sample type (sediment or water) and pond type (main pond, surface water pond, gravel pond, experimental pond) were both significant factors in determining community composition, and there was often a significant interaction between the two factors (see Table 2.2). Sample type explained a much greater proportion of the variation in community composition in prokaryotic communities (over 30 %) than in the other four communities sequenced, where the variance explained by sample type was always less than 10 %. For all five communities, the proportion of total variation in community composition explained by pond type and the sample type:pond type interaction was always less than 13 %.

Transforming the read abundance weighted dataset into a presence-absence dataset did not alter the overall patterns substantially. Some patterns appeared more pronounced, with R2 values increased compared to the read abundance datasets. Conversely, the interaction factor was found to be not significant for some communities, whereas it had been in the read abundance-weighted dataset. This may be because the presence-absence method gives greater weight to rare taxa than in the read-abundance weighted dataset.

Does a merged water sample represent the community of a larger pond?

The main pond in Pinkhill Meadow is 0.56 ha in area. My experimental analysis compared the taxonomic composition and read abundance of organisms retrieved from a single merged water sample (60 ml from 17 locations around the perimeter = 1020 ml) to the composition and mean read abundance from 17 separate 940 ml water samples taken at the same locations.

For prokaryotic taxa (bacteria and archaea), the merged sample not only detected all the phyla which were detected in the separate samples, but the relative read abundance in the merged sample of the top two phyla (Gammaproteobacteria and Actinobacteria) was very similar to the mean read abundance from the separate samples (within 3 % in both cases). The percentage read abundance of other phyla was not estimated as accurately.
However, the same pattern was not found in larger organisms: animals and green plants and algae. In these cases, the merged sample returned taxa which weren't detected elsewhere in the separate samples, such as Apiales, or were detected in the merged sample in a much greater abundance, such as Bacillariophyceae. The merged sample often failed to detect taxa which were found in the separate samples, such as Bivalvia or Brassicaceae (see table S2.3 for the full comparison).

This finding indicates that subsampling and then combining smaller volumes over a large pond provides a good representation of bacterial and archaean taxa, but not of animals, green plants or algae. This is probably because there are thousands of individual prokaryote organisms per ml of pond water, so a smaller volume is required per location to capture most of the genetic diversity, compared to the lower eDNA copy number of larger, patchily distributed eukaryotes.

Many studies have found that increasing the volume of water filtered increases the yield of eDNA (Hunter et al. 2019, Govindarajan et al. 2022), and pooling subsamples decreases species capture (Sato et al. 2017), yet there are still commonly used protocols and field equipment which recommend filtering smaller volumes, combining them, and capturing the eDNA on a single filter for a single pond (NatureMetrics 2022). When targeting eukaryotic organisms with eDNA metabarcoding, we recommend the use of larger water volumes in total, spread over several separate water samples covering different pond microhabitats. Each sample should be filtered, extracted from and sequenced separately to increase eDNA capture.

A few studies have examined the relationship between body size or total biomass and eDNA capture, but often within a single taxonomic group or species (Yates et al. 2020, Elbrecht et al. 2017), yet few have considered these implications for multi-marker eDNA metabarcoding,

where the target taxa may be several orders of magnitude different in abundance and biomass.

Conclusions

Considerations for study design when using eDNA metabarcoding for ponds

My study highlights some important considerations which should be made when designing future eDNA metabarcoding studies of ponds.

Firstly, multi-taxa eDNA metabarcoding has huge potential for use in freshwater pond environments, particularly for describing hitherto unknown pond biodiversity from many more branches of the tree of life, and for biomonitoring of entire pond ecosystems. However, it is no replacement for current, visual-based biomonitoring of pond ecosystem health, with very low detection of traditional macrophyte and macroinvertebrate indicator taxa, and restricted ability to resolve taxa to species level. Therefore, to make eDNA metabarcoding suitable for pond biomonitoring going forward, three approaches could be taken:

1. Increase the overlap with traditional taxa (macroinvertebrates and macrophytes), by optimising primers to better amplify the target organisms, expanding reference databases with more sequences from the target organisms and improving the taxonomic assignment pipelines so they are more efficient;

2. Produce new biomonitoring metrics based upon organisms which are detected well by 16S, 18S or COI metabarcoding e.g., Chlorophyta, meiofaunal assemblages (Dembowksa et al. 2018);

3. Develop taxonomy-free metrics by comparing eDNA communities across known diversity or disturbance gradients of ponds (calculated, for instance, by PSYM or the HSI for Great Crested Newts). There has been success using this method for other environments e.g., seabeds around Salmon aquaculture farms (Fruhe et al. 2020).

Secondly, at least for higher plants, and most probably for other organisms as well, eDNA sampled from ponds detects a large proportion of terrestrial organisms from the surrounding landscape. The pattern is more pronounced in water samples, whereas eDNA sampled from pond sediments detects fewer terrestrial organisms. This is probably due to ponds' small size relative to lakes, and low flow conditions, allowing eDNA to accumulate. The terrestrial 'contamination' may be a drawback for some applications, such as monitoring of aquatic pond species, but conversely could be an advantage for other research questions, for instance, monitoring change in the surrounding habitat or landscape, and examining the land-water interface.

Thirdly, eDNA is distributed heterogeneously throughout a pond, both between the pond water and the benthic sediments, and in the pond water throughout the area of the pond. For all five different communities sequenced by the five different taxonomic markers, water and sediment samples had a significantly different community composition (Figure 2.3, Table 2.2). Therefore, for a more complete picture of pond biodiversity, both pond water and surface sediments should be sampled in eDNA studies. If financial or time constraints limit sampling to either medium, we recommend sampling pond surface sediments over pond water: sediments do not require a filtration step, the extraction process usually is faster, and the yields of eDNA were, in our study, greater. Sediment traps could be used to record more accurately the timespan the eDNA was laid down.

In our larger pond (>0.5 ha), 12-16 separate one litre water samples detected a greater number of taxa than a single 1 litre sample of water merged from the 12-16 locations (Figure 2.5), other than for prokaryotic organisms. Therefore, we recommend that for ponds larger than 300 m^2, multiple water subsamples are collected, each of a sufficient volume (1 litre or greater) within financial and time constraints. If a pond has different microhabitats (e.g., reeds, bare substrate, submerged plants, open water, plant roots), care should be taken to sample each, as is also the case in the current sweep-net method (PSYM, Pond Action 2002).

Thirdly, my analyses suggest that beta diversity metrics based on eDNA are robust to different dataset pre-processing procedures. Whether taxa with low confidence of assignment were filtered out, and whether the dataset was read abundance-weighted or

purely based on the presence and absence of different taxa had very little impact on the relative effects of sample type (sediment or water) and pond type (gravel or clay substrate, pond location) on community composition. This pattern held true across all five communities sampled: prokaryotes, eukaryotes, fungi, animals and plants. The unfiltered analyses found a significant interaction between pond type and sample type whereas this was often lost in the filtered datasets. For biomonitoring multiple ponds, it may be better to use unfiltered datasets with no taxonomic assignment, as these better reflect the true genetic and community diversity of the samples (Cordier et al. 2020). Presence-absence data may be more appropriate for alpha diversity or gamma diversity metrics, but for beta diversity may give undue weight to rarer taxa (Machler et al. 2020).

Chapter 3: Understanding seasonal turnover in pond ecosystems using eDNA metabarcoding

Abstract

This chapter reports on the findings of whole-community eDNA metabarcoding at a high spatial and temporal resolution of ponds in a small wetland complex (~5 ha) in Oxfordshire, UK. Water and sediment samples were collected from 22 ponds in January, March, July, September and November 2020. The prokaryote, microbial eukaryote, multicellular eukaryote and macrophyte communities of the ponds were metabarcoded by amplifying eDNA from the samples using three different primer pairs (16S rRNA, 18S rRNA and the ITS2 gene region), followed by high-throughput sequencing on an Illumina MiSeq.

658 prokaryotic genera and 610 eukaryotic species were detected overall, a threefold increase in species detected using traditional visual-based identification methods. There were marked differences in community composition between water and sediment samples, however these were more distinct in prokaryote and microbial eukaryote communities than in mullticellular organisms. Sediment and water sample composition altered significantly between sampling timepoints for all four types of organism, but sediment community composition was more stable across seasons than community composition in water samples.

These findings have significant implications for environmental DNA metabarcoding methods in ponds going forward, indicating that sample season and water temperature should be considered in future sample design. Furthermore, this study sheds light on the community structure and diversity of microbes and microfauna in freshwater ponds, which have hitherto been little studied.

Introduction

Freshwater ponds (lentic waterbodies <5 ha in area and <5 m deep (Richardson et al. 2022)) are important habitats, both in the UK and globally, for biodiversity, ecosystem function and ecosystem services (Biggs et al. 2005, Céréghino et al. 2007, Céréghino et al. 2014). However, compared to rivers, lakes, and larger wetlands, they are under-researched (Biggs et al. 2017). In the UK and the rest of Europe, ponds are not included in the Water Framework Directive and there are only a handful of statutory instruments pertaining to their protection and monitoring compared to larger freshwater habitats (Hill et al. 2018, Oertli et al. 2010).

Environmental DNA metabarcoding - monitoring environments not through sampling and counting whole organisms, but through sampling their DNA - has now been deployed in a huge variety of ecosystems on all continents of the globe, from the Antarctic to the Tropics and from the deep ocean to the tops of mountains, although currently biased towards the Global North and West (Schenekar 2023). eDNA has been successfully extracted and sequenced from water, soil, sediment, air, and other media (Ruppert et al. 2019). In freshwater environments, eDNA is well on the way to becoming a standard monitoring tool. In ponds across England, a single-species eDNA assay is used to monitor Great Crested Newts (*Triturus cristatus)* (Natural England 2020) and garden ponds are monitored using eDNA via a citizen science programme (NHM 2022).

Due to research gaps, many unanswered questions about pond ecosystems remain. Traditional pond surveys, when they are been carried out, focus on macroinvertebrates and plants (macrophytes) (Biggs et al. 1998). eDNA metabarcoding in ponds has mainly focussed on amphibians, fish and macroinvertebrates (Harper et al. 2019a). Compared to rivers and lakes, the microbial, microfaunal and algal communities of ponds have rarely been studied, as have the ecological communities inhabiting pond sediments, not only the water column. Therefore, it is not known whether these communities are like those in other freshwater environments, such as rivers, lakes, and wetlands, or have unique characteristics. Additionally, ponds also have tended to be sampled in the summer months, and their community dynamics at other times of the year are less known (Hill et al. 2016).

These knowledge gaps are not only of interest for answering questions of fundamental ecology, but also of relevance to developing novel biomonitoring methods and metrics for ponds. For instance, if nationwide pond eDNA sampling is spread across several months, will the results be comparable? Does eDNA metabarcoding of microbes or microfauna

provide an insight into an aspect of overall pond ecosystem health and functioning hitherto overlooked?

eDNA metabarcoding studies of bacteria and archaea communities in sediment and water and in both rivers and lakes have found broadly consistent results: in water samples, Proteobacteria, Bacteroidetes, Actinobacteria and, to a lesser extent, Verrucomicrobia were the most abundant phyla (Cruaud et al. 2020, Li et al. 2020, Gweon et al. 2021, Doherty et al. 2017, Read et al. 2015), whereas in sediment samples, Proteobacteria and Bacteroidetes were still abundant, but Acidobacteria, Chloroflexi and Firmicutes predominated over Actinobacteria (Yuan et al. 2023, Liu et al. 2022, Wu et al. 2019, Liu et al. 2018, see Table S3.2).

eDNA metabarcoding of eukaryote communities in rivers and lakes have produced less consistent results, partly due to changes and inconsistencies in taxonomic classification (Adl et al. 2019). Phyla frequently reported from sequencing of eDNA in water and sediments include Ochrophyta, Arthropoda, Cryptophyta, Ciliophora, Chlorophyta and Dinoflagellata (see Table S3.1). However, there have been few studies which directly compare sequenced eukaryote communities from water and sediment from the same sample location. In Lake Baikal in Siberia, Yi et al. (2017) found similar proportions of the most abundant taxa in sediment and water samples (Metazoa, Chrysophyceae, Ciliophora, Cercozoa). Ionescu et al. (2022) used eDNA metabarcoding to sample both bacteria and eukaryote communities in ponds in north-eastern Germany, and found that 15 % of the variation in communities could be explained by sample type (sediment or water). Eukaryotic phytoplankton had similar taxonomic composition in water and sediments, with Chlorophyta the most abundant taxa. However, sediment samples had a greater proportion of Cryptophyta.

The ecological communities in freshwater lakes and ponds undergo seasonal changes. In higher latitudes, macrophyte abundance and diversity peaks in summer (Tian et al. 2023), and macroinvertebrate diversity in autumn (Hill et al. 2016). Phytoplankton may have “bloom” in spring, autumn or both (Winder & Cloern 2010), and zooplankton are also subject to seasonal variation (Sommer et al. 2012, Liu et al. 2022). Studies have shown that eDNA metabarcoding can reflect these seasonal changes in community composition. Metabarcoding of eukaryote communities has detected plankton blooms and die-offs in rivers (Cruaud et al. 2019) and lakes (Zhang et al. 2019, Banerji et al. 2018, Mikhailov et al. 2019, 2021). eDNA metabarcoding has also detected seasonal changes in bacteria composition in rivers (Liu et al. 2018), and lakes (Jiao et al. 2021) which appears to be more pronounced in water than in sediment samples.

Metabarcoding studies metazoan communities in freshwater environments have also found seasonal variation in community composition. For instance, in chironomid midges (Bista et al. 2017) and zooplankton from sequencing lake water (Yang and Zhang 2020), invertebrates in water samples from Sicilian streams (Hupalo et al. 2022) and fungal communities in stream water (Matsuoka et al. 2021).

eDNA transport and degradation is influenced by abiotic conditions, such as UV light, temperature, and wind-mixing (Barnes & Turner 2016), which also alter seasonally. Seasonal eDNA metabarcoding studies must take care to ensure changes in abundance and diversity metrics reflect a biological reality, rather than methodological artefacts There have been fewer metabarcoding studies targeting aquatic and wetland plants in freshwaters compared to those targeting animals. Most plant metabarcoding applications have sequenced ancient DNA in sediment cores rather than contemporary communities (Reveret et al. 2023). However, some whole community metabarcoding surveys of macrophytes have been carried out, for instance in wetland sediments in Australia (Shackleton et al. 2019), lake sediments in Norway (Alsos et al. 2018) and lake water in Michigan, USA (Drummond et al. 2021) and river water (Ji et al. 2021). Single species eDNA assays of aquatic vegetation, for instance, those detecting invasive species, have found temporal variation in the abundance of eDNA (Doi et al. 2021). I expect this finding to be repeated for the whole macrophyte community as it undergoes dramatic annual changes in abundance (Luo et al. 2016).

Few studies have directly compared the relative effects of sample medium (water, sediment, biofilm etc), sample season, and local abiotic conditions on communities sequenced via eDNA metabarcoding. This study is one of the first to do so (although see Ionescu et al. 2022). It is also the first, to my knowledge, to sequence whole green plant communities over the course of a year in any freshwater environment.

In this chapter, I report the results of a study monitoring four distinct components of the pond biological community across all seasons of the year, using eDNA metabarcoding methods. Water and sediment samples were taken at high resolution from 20 ponds in a small (5 ha) wetland complex (Pinkhill Meadow) in Oxfordshire, UK, along with water chemistry and other environmental variables, in January 2020. Sampling was repeated in March, July, September, and November 2020. Each of the c.320 samples generated from this sampling then underwent eDNA metabarcoding, via DNA extraction, amplification using three different primers and high throughput DNA sequencing of marker genes on an Illumina MiSeq. PCR primers selected to target different components of the pond community: bacteria and

archaea (16S rRNA gene), microbial eukaryotes (18S rRNA gene), multicellular eukaryotes (18S rRNA gene) and green algae and macrophytes (plant ITS2 gene region).

From observation and prior studies (Williams 2017, Williams et al. 2010, P. Williams personal communication Jan 2020), the environmental conditions of ponds at the study site varied considerably: both spatially, between different ponds, and temporally, throughout the course of a year, given varying rainfall, temperature and sunlight, and flood events from the nearby River Thames.

The aim of this study was threefold: to discern whether eDNA metabarcoding in sediment or water samples capture different aspects of pond ecological communities; to discern whether sample season influences community composition as sampled by eDNA metabarcoding; and to provide an insight on the taxonomic composition of microbial communities in ponds, to compare them with those in other freshwater habitats.

Hypotheses

Based on the research literature, the following hypotheses were formulated and tested:

1. Sediment and water samples will have a distinctly different taxonomic composition, reflecting taxa adapted to the different environmental conditions in pond sediments vs pond water (e.g., anoxic vs oxic, benthic vs pelagic).
2. Water samples will have greater seasonal variation in community composition than sediment samples. This is because the water environment undergoes more significant changes throughout the year in hydrology (water level, connectivity with other ponds) and other environmental variables (temperature, nutrient content) than the pond sediments, which have a more stable environment.
3. Pond physiochemical conditions (influenced by substrate, water source and shade levels) will also affect community composition.
4. Due to shorter generation times, higher numbers of individuals and higher taxonomic diversity, the patterns in hypotheses 1, 2 and 3 will be more pronounced in prokaryotic and eukaryotic microbes compared to multicellular plants, fungi and animals.

Methods

Study site

Pinkhill Meadow nature reserve, Oxfordshire, is a wetland complex of over 40 ponds, with varying sizes, shapes, depths, permanence and water sources (Figure 3.1). The ponds were excavated in 1990 to provide a practical case study for developing ideas on pond and wetland creation and management (Freshwater Habitats Trust 2019). The ponds were monitored, using traditional, visual identification-based methods, for wetland plants and invertebrates annually in the first decade, and then at intervals up until the present day (Biggs & Williams 2024).

The site is around 4 ha in size, and located in a meander on the River Thames, adjacent to Farmoor reservoir, to the west of the city of Oxford. Around 1.5 ha of the site is unimproved floodplain meadow and is grazed by a small herd of cattle between October and March and the hay cut once a year. Other than this, there is no management of the site, and it is not accessible to the public. The underlying geology of the site is a clay (alluvium) layer overlying a gravel aquifer. The alluvium layer is thicker at the northern end of the site, and so the ponds in this region have a clay substrate and tend to be semi-permanent, drying out in drought years. The ponds at the southern end of the site are dug into the gravel aquifer, and so are groundwater fed.

20 distinguishable water bodies were chosen in December 2019 for sampling for water, sediment and abiotic variables year-round. The 21 ponds on the site were divided into four rough "pond types" given similarities in substrate (clay vs gravel), physical habitat (vegetated vs open), hydrological regime (groundwater vs surface water-fed), and location on the site (see Chapter 2 methods for more details). Figure 3.1 shows a map of the site with locations of sample points named in red. Some larger waterbodies were sampled in multiple locations around the perimeter of the pond. Smaller waterbodies were sampled in a single location near the perimeter.

Full sampling of water and sediment and in-field testing of abiotic variables over all 21 waterbodies was first carried out over a single day in January 2020. The full sampling was then repeated bi-monthly over a single day in March, June (sampling was delayed from May

due to restrictions caused by the Covid-19 pandemic), July, September and November 2020. For full details of water and sediment samples taken from each pond per sampling event, along with blanks, see supplementary Table S3.5.

Water samples were collected from 15 waterbodies (6 samples were collected from the Main Pond, and a single sample from 14 other ponds) and analysed in a laboratory for additional water chemistry variables. These samples were taken in June, September and November 2020, as close as possible to the main sampling event (either the prior or subsequent day). Not all sampling points were sampled at every sampling event due to changing water levels – in some months, ponds did not have any water in them. In other events, separate ponds were joined due to floodwater. A maximum of 42 water and sediment samples were collected from the 21 waterbodies on site. The samples from June were used for methods development and testing, and the results are reported in Chapter 2.

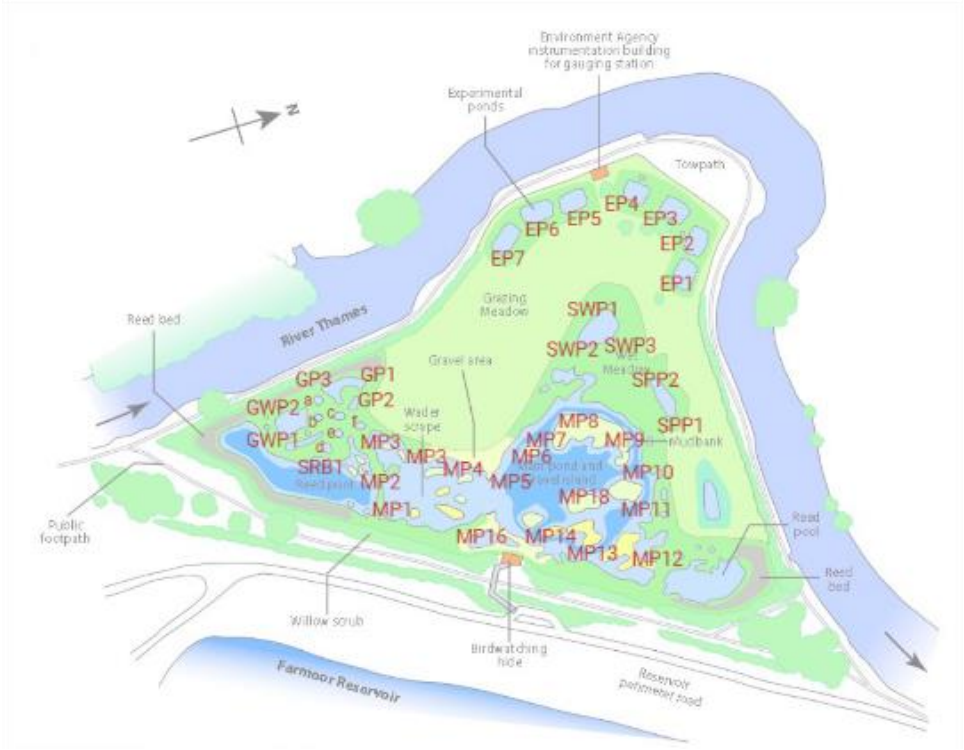


Figure 3.1: Pinkhill Meadows pond complex site, with key human and natural features labelled. All 42 sample points overlaid in red text. MP = Main Pond, SPP = Semi-permanent pond, SWP = Surface water pond, EP = Experimental pond, GWP = Groundwater pond, GP = Gravel pond, SRB = Southern reed bed.

Water and sediment sample collection

Water samples were collected using 1L plastic bottles which had been sterilised via a laboratory acid washer or soaking in 10 % bleach. Nitrile gloves were used to collect water and changed between samples. Bottles were triple washed in pond water before a sample was taken. Water samples were taken around 10cm from the surface of the pond in all cases (a compromise between the tendency of DNA to sink and avoiding disturbing the sediment).

Water was stored in coolboxes and transported to laboratory facilities within four hours of collection, where it was stored at 4°C for a maximum of 24h before filtration. Water samples were taken from the same point in all sampling events.

Sediment sample collection and storage

Sediment samples were collected from the same location as the water samples, using a custom sediment sampler comprising of a tubular aluminium rod with a holder for a 50 ml centrifuge tube, both of which had been sterilised using an acid wash or 10 % bleach before use. The scoop was triple washed in pond water before each sample. Sediment samples were taken from the surface of the sediment, within the top ~10cm.

To attempt to prevent cross contamination between sample sites, the sediment scooper was wiped with 10 % bleach between uses and left to air dry. Sediment samples were stored in coolboxes and transported to the laboratory within 4 hours of collection and stored at -20’C.

Measuring abiotic variables

In the field, an ultrameter II was used to measure pH, temperature, Total Dissolved Solids (TDS), Conductivity and Oxidative-Reductive Potential (ORP) at each sample point. Measurements were taken at the same time as the water and sediment samples were collected. The sensor was triple washed in pond water before the measurements were taken. For each sample point, three measurements were taken, and the mean calculated.

Water depth at each sample point was measured using a metre rule.

Water samples for water chemistry analysis were collected using a wide-mouthed bucket which was single washed before sampling. From this larger sample, a smaller 100 ml sample was taken to quantify total suspended solids (TSS) in the laboratory. Two 60 ml samples were subsampled from the bucket and filtered using 0.45um filter on site. The bucket was then stirred, and a third 60 ml sample was collected and not filtered. The two

filtered samples were analysed for 1. total Soluble Reactive Phosphorus (SRP) and 2. chlorophyll, ammonia (NH_4), Total Dissolved Nitrogen (TDN) and Total Organic Carbon (TOC). The unfiltered sample was analysed for Total Phosphorus (TP). For details of the full method, see Bowes et al., (2018).

CanopyApp (University of New Hampshire, ver. 1.0.4) was used to estimate tree canopy cover at each sample point in July 2020. The site was visually assessed during each sampling event and photographs taken to record which waterbodies were hydrologically connected or separated due to flooding or drying up.

Filtration

An average of 1000 ml of water was collected for eDNA filtration at each sample point. For 14 smaller ponds, only one 1000 ml sample was taken. For three medium ponds (>100 m^2, GWP, SPP and GP1), two 1000 ml water samples were taken in the field. In the laboratory, each of these were sub-sampled to 500 ml which were combined to make one 1000 ml sample which was subsequently well mixed and filtered.

One medium-size pond (SWP) was sampled in a similar way, except three 1000 ml samples were taken in the field, and these were sub-sampled for 330 ml each which were combined to make one 990 ml sample. For the large main pond, up to 17 separate 1000 ml samples were collected in plastic bottles in the field. In the laboratory, ~60 ml of each of these samples were subsampled and combined to make a merged Main Pond sample of ~1000 ml, which was subsequently filtered. The ~940 ml of the 17 separate samples (MP1 – MP17) were also filtered separately.

All samples were filtered in a sterile environment using vacuum pump filters sterilised via acid washing or soaking in 10 % bleach. Each 1000 ml sample was initially pre-filtered with a 12um Cellulose-Nitrate filter (Whatman, AE100) and subsequently filtered with a 0.45um Cellulose Nitrate filter (7141 114). The filter papers were removed from the filters using sterilised tweezers and placed in 5 ml centrifuge tubes, which were stored at 4°C until extraction.

DNA extraction

DNA from water samples was extracted from the stored filter papers after defrosting using the standard protocol of the DNeasy PowerWater DNA extraction kit (Qiagen Group). An extraction blank was produced by following the normal extraction protocol but omitting any

sample. Concentration of DNA was determined using a Nanodrop spectrophotometer. DNA was stored in Elution buffer in 1.5 ml microcentrifuge tubes at 20°C. The total volume of extracted DNA ranged from 70 µl – 100 µl.

DNA from sediment samples was extracted by defrosting the full sediment sample and then subsampling 0.25g, using a clean weighing boat and disinfected tweezers. DNA was then extracted using a DNeasy Powersoil kit (Qiagen Group) following manufacturers' instructions. An extraction blank was produced by following the usual protocol but omitting any sample. Concentration of DNA was determined using a Nanodrop spectrophotometer. DNA was stored in Elution buffer in 1.5 ml microcentrifuge tubes at –20'C. The total volume of extracted DNA ranged from 80 µl – 85 µl. All DNA extraction protocols were carried out using sterile procedures i.e. with gloves, and the use of 99 % ethanol and bleach to sterilise equipment and laboratory benches between any handling of DNA samples.

Amplification and sequencing

Three primer pairs were chosen, which amplify different DNA or RNA targets and produce different length fragments, in order to detect different segments of the pond community (see Table 3.1 below).

Table 3.1: Primer pairs used in the Pinkhill Meadow seasonal study. For full primer sequences, see Table S2.1

Primer pair name and reference	**Fragment length**	**Gene amplified**	**Community primer optimised to target**
NSF573/EKNSR951 (Mangot et al. 2013)	425bp	18S rRNA	Eukaryotes
515F/806R (Caporaso et al. 2011 and Walters et al. 2015)	390bp	16s rRNA	Bacteria
ITS2-S2F/ITS4_R (Chen et al. 2010)	300-460bp	ITS2 rDNA	Viridiplantae (vascular plants and green algae)

DNA was amplified using a two-step PCR approach. Each primer pair was first tested on six randomly chosen samples (three water and three sediment), along with an extraction blank and PCR blank, to assure positive amplification. Then, all samples were amplified with a modified primer (an amplicon primer with Illumina MiSeq sequencing primer and pre-adaptor added). 1-2 PCR blanks were also amplified at this stage (normal PCR reagents, but with molecular grade water added rather than any sample). Then Step 2 PCR was carried out to

add on the barcodes (Illumina MiSeq index) and flow-cell adaptors. Steps 1 and 2 were repeated for the two other primer pairs. Therefore, each sample had three two-step PCRs carried out on it, one per primer pair. For PCR conditions, see Table S3.5.

Amplicons were normalised using SequalPrep Normalisation Plate Kit, 96-well (Invitrogen, Carlsbad, CA), gel purified using the QIAquick gel extraction kit (Qiagen Group) and quantified using Qubit high sensitivity dsDNA Assay kit (Invitrogen, Carlsbad, CA). The resultant amplicon library was sequenced at a concentration of 9 pM with a 0.675 pM addition of an Illumina generated PhiX control library. Sequencing was performed on an Illumina MiSeq platform using MiSeq Reagent Kit v3 (Illumina Inc., San Diego, USA).

Bioinformatic pipeline

For the 16S and 18S DNA sequences, raw reads were processed through the DADA2 pipeline ver. 1.8 (Callahan et al., 2016) in R (R Core Team, 2018). Briefly, adapters and primers were initially removed from the raw reads using cutadapt (Martin 2011). Then amplicon reads were trimmed to maintain Q score > 30, at 230 and 200 bases, forward and reverse respectively for 16S reads and at 250 and 200 bases respectively for 18S. The results filtered with DADA2 default settings, except for the maximum number of Ns (maxN) = 0 and maximum number of expected errors (maxEE) = c(5,5).

Sequences were dereplicated and the DADA2 core sequence variant inference algorithm applied. Processed forward and reverse sequences were merged using the mergePairs function, and a sequence table was constructed from the resultant, merged amplicon sequence variants (ASVs). Chimeric sequences were removed from the ASV table using remove BimeraDenovo with default settings. To assign taxonomy, the online IDTAXA tool from the R package DECIPHER ver. 2.14 (Murali et al. 2018) was used, using the SILVA database, ver. 138.1 (10th March 2021) for 16S and the PR2 database, version 4.13 (17th March 2021), both provided by the DADA2 package. This function gives a taxonomic classification to genus level and provides confidence scores for each level of classification. The confidence threshold for both types of sequence was set at 60 %.

For 16S ASVs, using the R package ‘phyloseq’ (version 1.42.0, McMurdie & Homes 2013) samples of less than 4,000 reads were pruned, leaving 288 samples. Ultra-rare taxa (less than three reads and only in one sample) were removed. Samples were rarefied to the median sampling depth (21,605 reads). Removing low-read samples and taxa aids in the removal of false positive detections (Shirazi et al. 2021, García-Machado et al. 2023). Rarefaction is a widely used and statistically valid way to normalize sample size, and so

control for the uneven sequencing effort between samples (Weiss et al. 2017, Schloss 2023).

ASVs with the same taxonomic assignment were agglomerated, and those with a root confidence of <100 were excluded from the dataset. ASVs assigned to ‘Mitochondria’ and ‘Chloroplast’ were also removed. Any field, extraction or PCR blanks that had not been removed from preceding steps were then removed.

16S blanks contained 123 taxa, of which all but two were shared with pond water and sediment samples. However, only five taxa had total read abundance across all blanks >1000. These five taxa were removed from further analysis.

For 18S ASVs, a similar procedure was followed, although samples of less than 1,000 reads were pruned, leaving 285 samples. As before, ultra-rare taxa were removed, the samples were rarefied to the median sampling depth (8,908 reads), ASVs were taxonomically agglomerated and those of root confidence <100 were excluded. To separate the datasets for microeukaryotes and multicellular eukaryotes, taxa assigned as “Metazoa”, “Streptophyta” and “Fungi” were filtered out from the dataset for the former and kept for the latter (with all other taxa filtered out). Any field, extraction and PCR blanks that had not been removed from preceding steps were then removed. Blanks contained 32 taxa, all of which were also present in environmental samples. Only one taxon was present at over 1000 reads across all blanks, which was identified as a Vermamoebidae Amoebezoan, and was removed from subsequent analyses.

For the ITS2 sequences, a different process was followed. Raw reads were processed and taxonomically assigned using the HONEYPI pipeline implemented in Python 2.7 (Oliver et al. 2021). The HONEYPI pipeline removes adaptors and quality filters the raw reads using TrimGalore v.0.6.4, and then uses the DADA2 pipeline to generate an Amplicon Sequence Variant (ASV) abundance table containing chimera-removed, high-quality error-corrected sequences. For each ASV, conserved regions flanking ITS2 are removed with ITSx v.1.1b; and resulting sequences taxonomically classified using the naive Bayesian classifier against an in-house ITS2 database of 966,676 (25th March 2020). Unless stated otherwise, default parameters were used for the steps listed.

Each of the ASVs was taxonomically assigned to species level by the RDP classifier within the HONEYPI pipeline and given a confidence level of assignment between 0 and 1.

Using the R package ‘phyloseq’ samples of less than 4,000 reads were pruned and ultra-rare taxa were removed as before. Samples were rarefied to the median sampling depth (19,766 reads). ASVs with the same taxonomic assignment at species level were agglomerated. Taxa assigned as “NA”, “Fungi”, “SAR”, “Amoebozoa”, “Straminipila” and “Metazoa” were filtered from the dataset. Any field and PCR blanks that had not been removed from preceding steps were then removed.

The pITS blanks contained 48 taxa, all of which were also found in pond water and sediment samples. Six of these taxa, *Alium cepa, Raphanus sativus, Juglans regia, Fraxinus excelsior, Lythrum salicaria* and *Phagrmites australis* were present at high read numbers in blank samples (>1000 reads over all blanks). However, the latter four were present at superabundant levels in environmental samples (>73,000 reads for each taxon), indicating contamination from pond samples to field blanks. The former three could be due to contamination in the laboratory stages, but could also have originated environmentally, so were not removed.

Statistical analyses

All statistical analyses were carried out in R, version 4.2.0 (R Core Team 2021). To assess differences in abiotic variables between months and pond types, ANOVAs and Kruskal-Wallis tests were carried out using functions in base R. Abiotic variable plots were made by using the trans_diff, cal_diff and plot_diff functions in the R package microeco, version 0.11.0 (Liu et al. 2021).

Almost all other statistical analyses were carried out in the R package microeco, version 0.11.0 (Liu et al. 2021). Venn diagrams were drawn using the function plot_venn in the trans_venn class. Relative abundance of different taxa was calculated, and bar graphs plotted, using functions within the trans_abund class. Taxa abundance between different sample types, pond types and months was compared via t.test and Kruskal-Wallis test using functions within the trans_diff class. Three measures of alpha diversity (Simpson, Shannon and Chao) were calculated and compared via t.test or Wilcoxon test within the trans_alpha class.

To examine beta diversity, functions within the trans_beta class were used. Beta diversity was calculated using a Bray-Curtis index and a Principle Co-ordinates Analysis was performed using the cal_ordination function. A permutational multivariate analysis of variance (PERMANOVA) was performed using the cal_manova function, which uses the adonis2 function in the vegan package, version 2.6-4 (Okansen et al. 2022). Initially, this was

performed on all samples using only sample type (sediment vs water) as a factor to examine the effect of sample type on community composition. Then, to examine the effect of season and pond type on community composition, sediment and water samples were separated and PERMANOVA performed, using sample month, pond type and their interaction as factors.

All higher plant (Embryophyta) taxa in the pITS dataset (n = 46) were assigned a 'wetland status' function according to the species list used in the National Pond Survey (Biggs et al. 1998). Floating-leaved and submerged species were combined into an "aquatic" category, whereas emergent species and trees or shrubs were combined into a "wetland" category. Plants which did not appear on the NPS list were classified as "terrestrial".

Results

eDNA concentrations, reads and ASVs

After filtration and extraction, the concentrations of eDNA from pond water samples ranged from 0.76 µg/µl to 114.1 µg/µl (*M* = 25.62 µg/µl, *SD* = 24.12 µg/µl). In sediment samples, post-extraction the concentration ranged from 2.32 µg/µl to 168.2 µg/µl (*M* = 60.03 µg/µl, *SD* = 2.5 µg/µl).

After processing with the DADA2 pipeline (quality trimming, de-replication, chimera removal), the raw datasets were as follows: for 16S sequences, this process produced 102,391 ASVs, and the total number of reads across 316 samples was 7,548,590; for 18S, 3,483,675 merged sequences were recorded in 14,531 ASVs and 308 samples; and for ITS2 sequences, 11,351,976 raw reads across 8,642 ASVs and 326 samples.

After rarefaction, removal of low-read samples and ultra-rare taxa, removal of taxa with low confidence of taxonomic assignment and taxonomic agglomeration (pooling ASVs with the same taxonomic assignment together), the final datasets contained the following number of samples and ASVs: the 16S dataset contained 658 taxa across 280 samples (120 water and 160 sediment); the 18S microeukaryote dataset contained 268 taxa across 281 samples (125 water and 156 sediment); the 18S multicellular dataset consisted of 120 taxa across 280 samples (124 water and 156 sediment. Finally, the plant ITS2 (pITS) dataset contained 222 taxa across 323 samples (143 water and 180 sediment.)

Physiochemical conditions of the Pinkhill Meadow ponds

The River Thames water level, the average rainfall and the temperature varied widely throughout the sampling period January – November 2020 (see Fig 2). Therefore, the hydrological connectivity of different ponds in the complex also varied greatly (see table S6 and images S1-5). As anticipated, pond depth differed significantly between different sampling events (ANOVA, $F(4) = 23.47$, $p < 0.001$), with the mean depth in July and September significantly lower than the mean depth in January and November (Tukey HSD test, all comparisons $p\ adj < 0.001$).

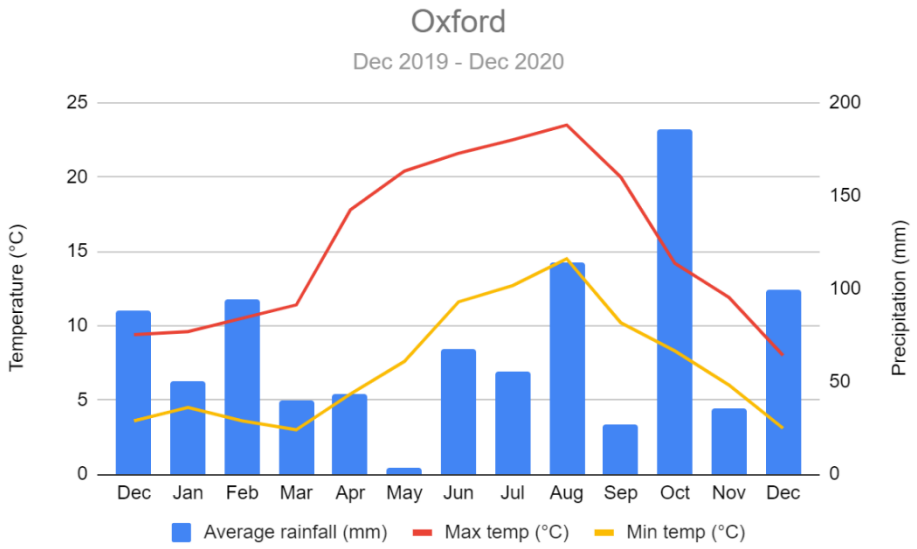


Figure 3.2: Data for climate variables at Oxford weather station throughout the sampling period. Data provided by the Met Office and extracted from https://www.metoffice.gov.uk/research/climate/maps-and-data/historic-station-data *on 9th February 2022*

Percentage canopy cover results from June and September were combined and compared between pond types to give an indication of differing shade levels across the study site. Canopy cover did not differ significantly between pond types (Kruskal-Wallis test, $H(3) = 3.67$, $p = 0.299$). However, from observation the gravel-substrate ponds in the south of the site and the experimental ponds tended to be more wooded and have a higher canopy cover.

Water chemistry conditions varied markedly between different sampling events (see Fig 3). Analyses of variance for each variable by month were carried out, and all were found to be

significant (all $p < 0.001$, see Table S3.7). For most variables, all months were found to be significantly different from each other.

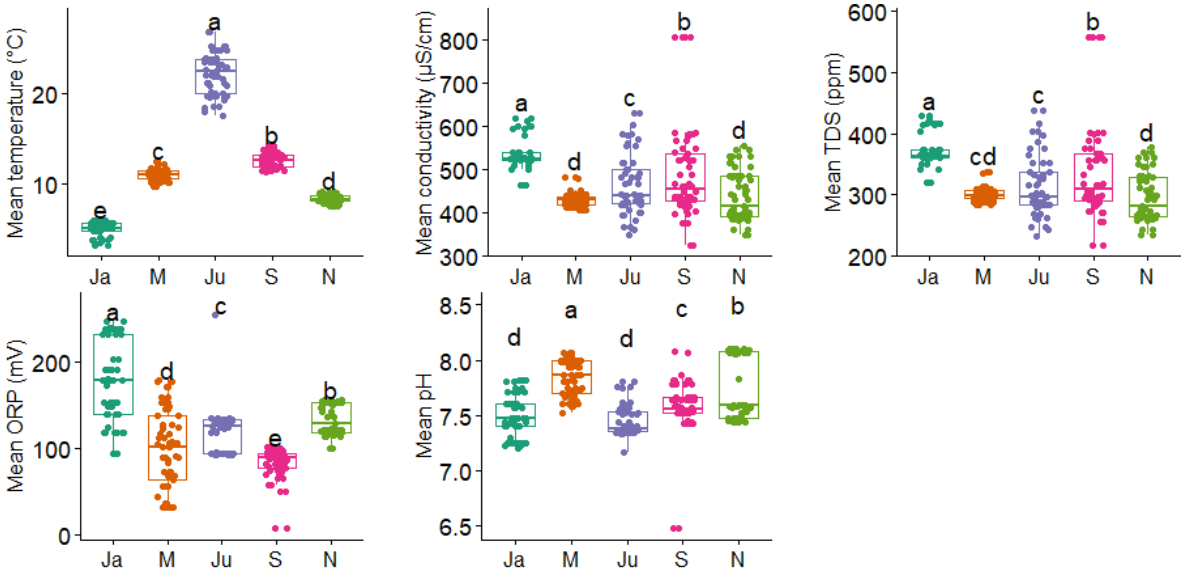


Figure 3.3: One-way ANOVA results as a boxplot for each physiochemical variable by month. p value for all ANOVAs was <0.001. Sample size n = 185. Lower case letters indicate if individual months' means are significantly different, post-hoc Tukey HSD tests at the p adj = 0.05 level.

Due to funding constraints, water samples for laboratory chemical analysis were only collected during the September and November sampling events. Mean values from September and November were compared using multiple Wilcoxon tests with a Bonferonni-Holm correction, and no significant differences in mean values between these months were found for any of the chemical variables analysed.

To compare conditions between pond types, Kruskal-Wallis analyses were performed on the ultrameter and laboratory datasets collected. Eight variables varied significantly ($p < 0.05$) between pond type: soluble reactive phosphorus, ammonia, total dissolved nitrogen, dissolved organic carbon, chlorophyll-a, pH, total dissolved solids and conductivity (see Figure 3.4 and table S3.8 for full statistical reporting).

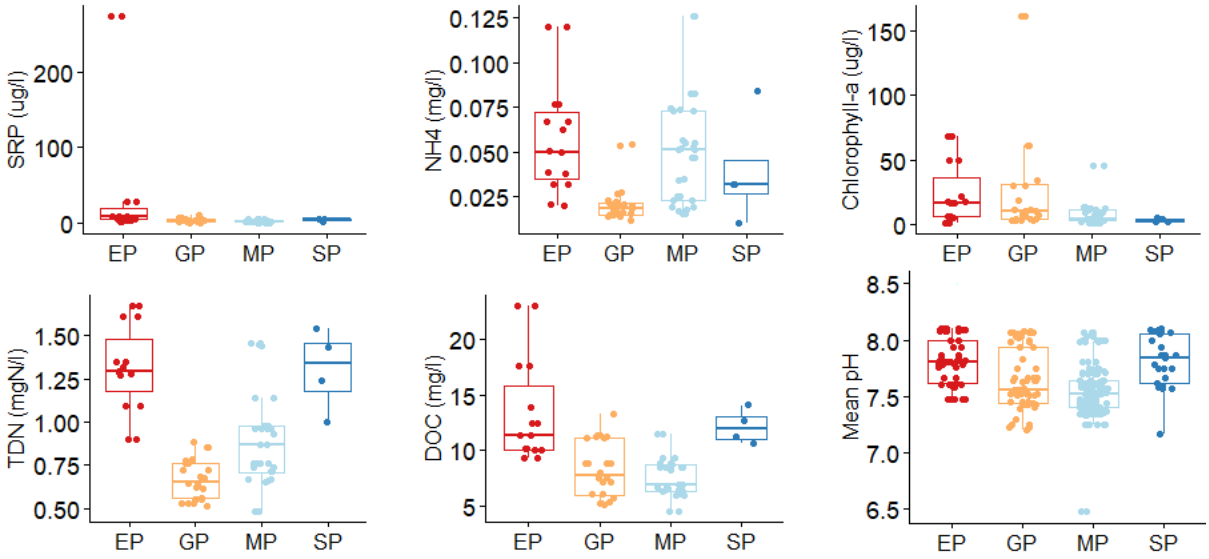
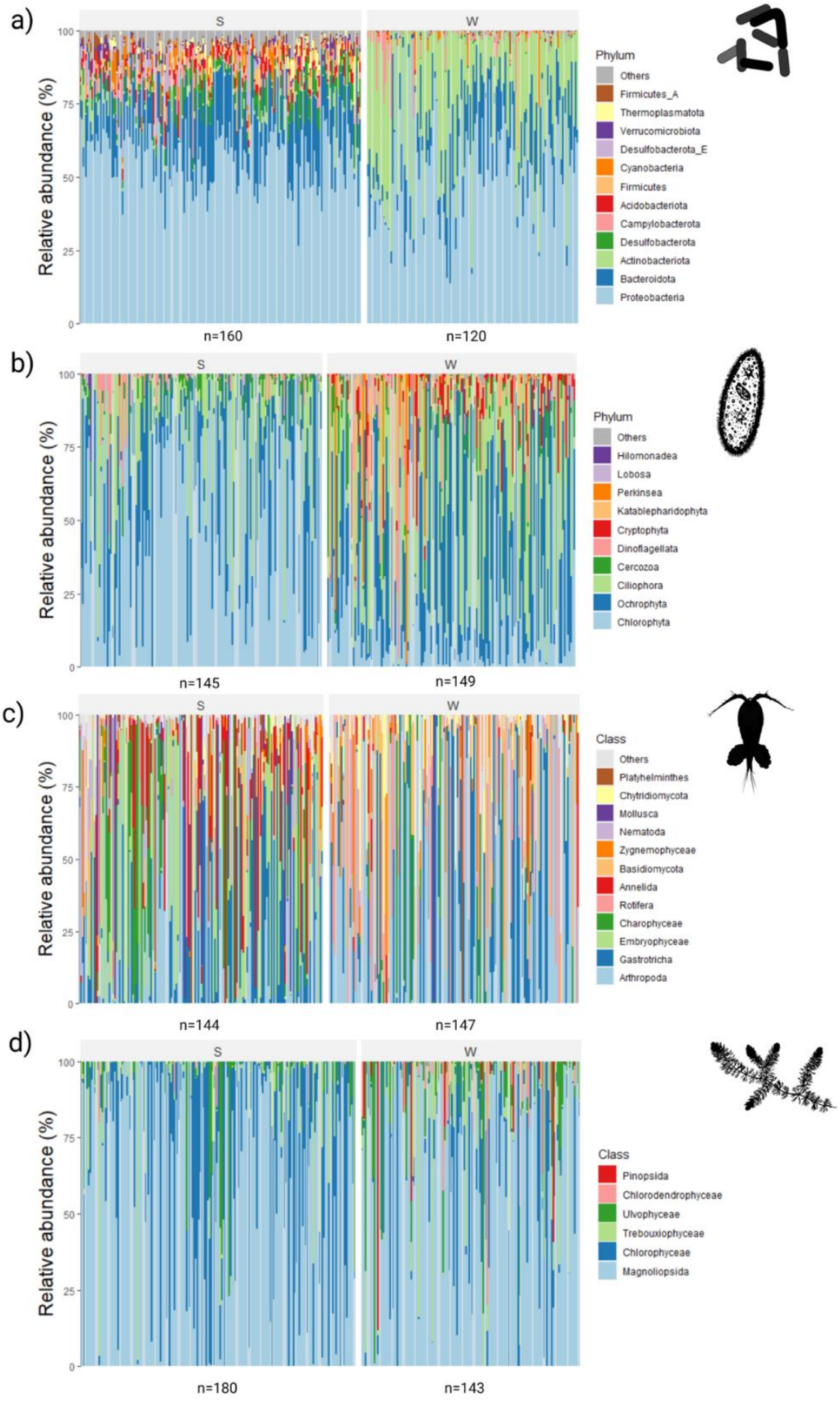


Figure 3.4: Kruskal-Wallis one way analysis of variance test results as a boxplot for each physiochemical variable by month. p value for SRP and NH_4 <0.01. p value for other measurements <0.001. Sample size n=185. "SRP" = soluble reactive phosphorus, "TDN" = total dissolved nitrogen, "DOC" = dissolved organic carbon

Taxonomic composition of the Pinkhill Meadow ponds

Figure 3.5 (below): Taxonomic composition of a) bacteria and archaea b) microbial eukaryotes c) multicellular eukaryotes and d) green plants and green algae of Pinkhill Meadow ponds, across all 5 sampling events using metabarcoding. 'Relative abundance (%)' = relative read abundance, 'S'= Sediment samples, 'W'=Water samples.


a)
S
W
Relative abundance (%)
100
75
50
25
0
n=160
n=120
Phylum
Others
Firmicutes_A
Thermoplasmatota
Verrucomicrobiota
Desulfobacterota_E
Cyanobacteria
Firmicutes
Acidobacteriota
Campylobacterota
Desulfobacterota
Actinobacteriota
Bacteroidota
Proteobacteria
b)
S
W
Relative abundance (%)
n=145
n=149
Phylum
Others
Hilomonadea
Lobosa
Perkinsea
Katablepharidophyta
Cryptophyta
Dinoflagellata
Cercozoa
Ciliophora
Ochrophyta
Chlorophyta
c)
S
W
Relative abundance (%)
n=144
n=147
Class
Others
Platyhelminthes
Chytridiomycota
Mollusca
Nematoda
Zygnemophyceae
Basidiomycota
Annelida
Rotifera
Charophyceae
Embryophyceae
Gastrotricha
Arthropoda
d)
S
W
Relative abundance (%)
n=180
n=143
Class
Pinopsida
Chlorodendrophyceae
Ulvophyceae
Trebouxiophyceae
Chlorophyceae
Magnoliopsida

Across all ponds and all timepoints, the most common Bacteria phyla by read abundance were Proteobacteria (sediment: *M* = 57.6 %, *SD* = 10.4 %, water: *M* = 53.4 %, *SD* = 18.2 %) and Bacteroidota (sediment: *M* = 17.9 %, *SD* = 8.7 %, water: *M* = 14.1 %, *SD* = 10.9 %). In water samples, Actinobacteria made up most of the remainder (*M* = 28.2 %, *SD* = 18.3 %), whereas in sediment samples Actinobacteria composed on average only 1.82 % of the total (*SD* = 3.4 %). Desulfobacterota and Desulfobacterota_E were also relatively common in sediment samples, on average 4.0 % (*SD* = 3.5 %) and 2.0 % (*SD* = 0.5 %) of the total reads respectively. Acidobacteria and Firmicutes were present in sediment samples at low read abundances (*M* = 2.7 %, *SD* = 2.0 % and *M* = 2.2 %, *SD* = 2.6 % respectively), whereas these phyla were virtually absent from water samples (see Figure 3.5a). All the 30 prokaryote phyla present exhibited significant differences in read abundance between sediment and water samples (multiple *t-tests* with fdr p value adjustment, all *p adj* < 0.05).

The top 20 most abundant genera in both water and sediment samples made up over 50 % of total read abundance by month (see Figure S3.1, a and b). Often, these genera have been described only recently and there is only one type species. The level of confidence of taxonomic assignment was variable. In water samples, the top 5 most abundant genera were assigned to *Rhodoluna, Sphingorhabdus_B, Planktophila, Aquirickettsiella* and *UBA952.* In sediment samples, the top 5 most abundant genera were assigned to *LD21, SCTMO1*, *Azonexus, Halioglobus* and *UBA4417.*

At the phylum level, the microbial eukaryote community was composed mainly of Chlorophyta (sediment: *M* = 56.0 % of reads, *SD* = 30.2 %, water: *M* = 19.8 %, *SD* = 22.8 %), Ochrophyta (sediment: *M* = 23.5 %, *SD* = 22.6 % and water: *M* = 31.7 %, *SD* = 26.2 %) and Ciliophora (sediment: *M* = 14.6 %, *SD* = 19.6 %, water: *M* = 26.3 %, *SD* = 24.6 %). Water samples also had notable levels of Cercozoa (*M* = 4.8 %, *SD* = 9.7 %), Cryptophyta (*M* = 4.9 %, *SD* = 8.7 %), Katablepharidophyta (*M* = 3.9 %, *SD* = 9.8 %) and Dinoflagellata (*M* = 3.6 %, *SD* = 10 %) reads (see Figure 3.5b). When comparing read abundances of phyla in sediment and water samples, Chlorophyta were significantly more abundant in sediment samples, whereas Ochrophyta, Ciliophora, Cercozoa, Cryptophyta, Katablepharidophyta and Perkinsea were higher in water samples (multiple *t-tests* with fdr p value adjustment, all *p adj* <0.05).

In water samples, the top 25 genera made up around 75 % of the total read abundance (see Figure S3.2a). The 5 most abundant genera were *Limnostrombidium, Navicula,* an unidentified member of the Crustomastigaceae family, and unidentified members of *Synurales* and Katablepharidales orders. In sediment samples, the top 25 genera by read

abundance made up only around 50 % of the total. The 5 most abundant genera were assigned to *Scotinosphaera, Navicula, Desmodesmus, Loxodes* and *Stentor* (Figure S3.2b).

In the multicellular eukaryote community, Arthropoda (sediment: *M* = 14.7 %, *SD* = 25.8 %, water: *M* = 31.6 %, *SD* = 37.5 %) and Gastrotricha (sediment: *M* = 19.1 %, *SD* = 25.4 %, water: *M* = 22.0 %, *SD* = 30.2 %) were dominant classes. Charophyceae (Stoneworts, traditional indicators of a clean water habitat) made up on average 11.2 % (*SD* = 23.8 %) of reads in sediment and 4.6 % (*SD* = 16.7 %) of reads in water. Embryophyceae and Annelida were common in sediment samples, comprising a mean of 19.1 % (*SD* = 27.9 %) and 12.1 % (*SD* = 22.2 %) of reads respectively. In water samples, Rotifera were common, making up an average of 13.6 % (*SD* = 26.0 %) of reads, as were Basidiomycota, making up 8.8 % (*SD* = 19.7 %). Nine classes had significantly different read abundance between sediment and water samples (multiple *t-tests* with fdr p value adjustment, all *p adj* <0.05). Embryophyceae, Charophyceae, Annelida, Mollusca and Nematoda were all more abundant in sediment samples, whereas Arthropoda, Rotifera, Basidiomycota and Chytridiomycota were more read abundant in water samples.

Again, the 25 most abundant genera composed around 50 % of the overall read abundance in both sediment and water samples (see Fig S3 a and b). In water samples, the top 5 genera by read abundance were *Chaetonotus*, *Cyclops, Chara, Eucyclops* and *Macrocyclops.* In sediment samples, *Chara* and *Chaetonotus* were also highly abundant, joined by *Pisidium, Sparganium* and *Sminthurides.*

In the Viridiplantae community (green plants and algae) at the phylum level, both water and sediment samples had a similar taxonomic composition, with Streptophyta (macrophytes - 'higher' plants and charophytes) comprising around of 65 % of the reads (sediment: *M* = 64.1 %, *SD* = 35.5 %, water: *M* = 69.7 %, *SD* = 31.6 %) and Chlorophyta (green algae) comprising 35 % of the reads (sediment: *M* = 35.9 %, *SD* = 35.5 %, water: *M* = 30.3 %, *SD* = 31.6 %, figure 3.5d).

The top 25 species made up between 50 % and 75 % of the total read abundance in both sediment and water samples (Figure S3.4 a and b). The most abundant species in sediment samples were *Phragmites australis, Lythrum salicaria, Berula erecta, Filipendula ulmaria* and *Sparganium stoloniferum.* In water samples, *Berula erecta* was also amongst the most abundant species, in addition to *Fraxinus excelsior, Urtica dioica, Cucumis sativus* and *Planophila bipyrenoidosa.*

Of the 96 'higher plant' (Embryophyta) species identified, 61 taxa were terrestrial species, 33 were wetland species and only two were aquatic species. When examining read abundance of these different categories, water samples were composed of around 70 % reads from terrestrial species, much higher than the 30 % found in sediment samples (*t test, p adj* < 0.001). By contrast, sediment samples had a much higher proportion of wetland species' reads (68 %) than water samples' 23 % (*t test*, *p adj* < 0.001). Water samples had a marginally higher percentage of reads from aquatic species (*t test*, *p adj* = 0.023).

27 of the 96 Embryophyte species were cultivated plants, such as Allium cepa (onion), Brassica oleracea (cabbage) and Tricitum turgidum (durum wheat). These cultivated species made up 11.7 % of total read abundance in sediment samples, and 24.4 % of total read abundance in water samples.

Community dissimilarity between sediment and water samples is more pronounced in microbes than in multicellular organisms.

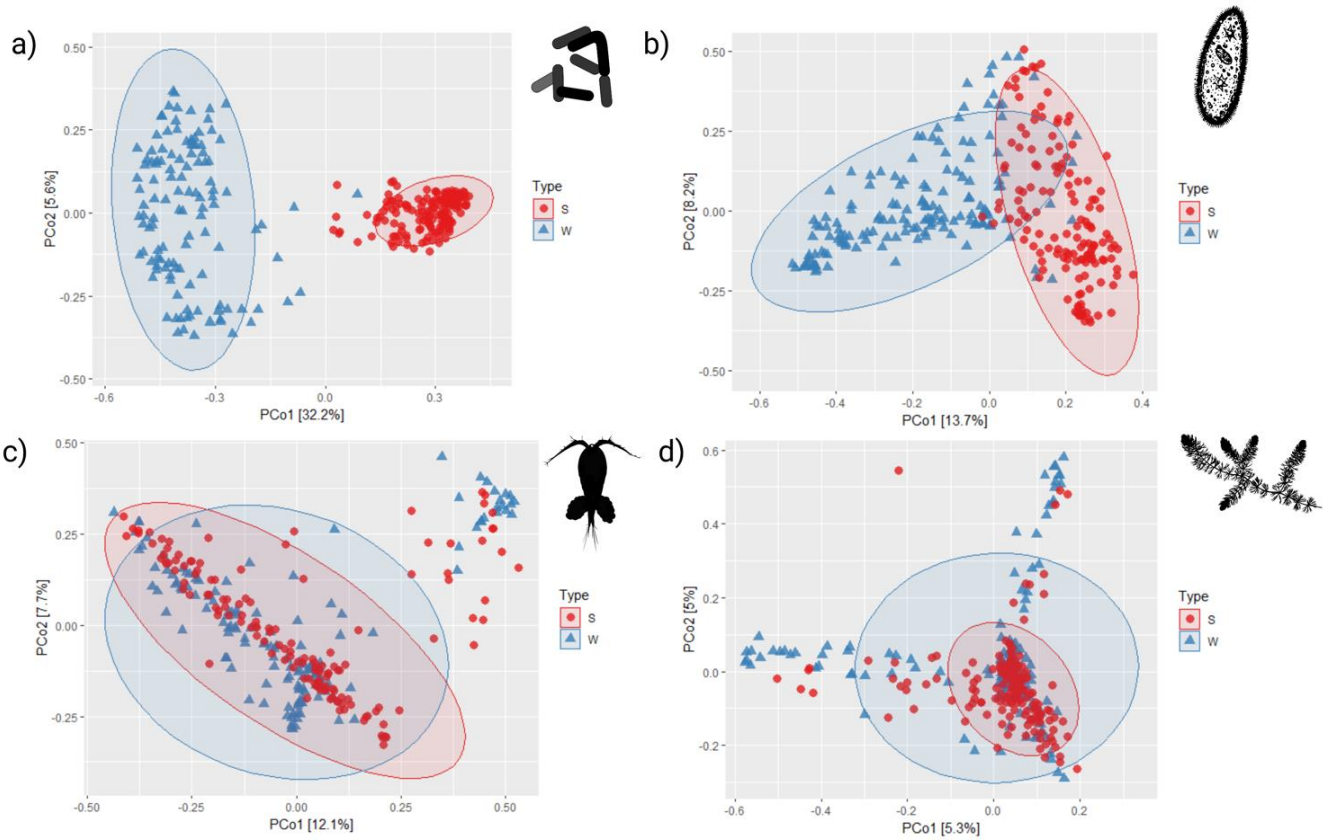


Figure 3.6: Principal co-ordinate analysis (PCoA) plots of Bray-Curtis distances between samples of different sample type (red = sediment samples, blue = water samples) in a) bacteria and archaea, b) microbial eukaryotes c) multicellular eukaryotes and d) green plants and green algae. Ellipses show statistical confidence at the 95 % level.

The final bacteria and archaea dataset contained 658 taxa across 280 samples (120 of these were water samples, and 160 sediment samples.) The majority of taxa (n = 347) were found in both types of sample, and these taxa accounted for 90.7 % of total read abundance. Sediment samples had 251 unique taxa and water samples had 60 unique taxa. In addition, sediment samples had a significantly higher mean alpha diversity than water samples on three measurements when compared by a *t test*: Simpson, Shannon and Chao alpha diversity (all *p adj* < 0.001, see Figure S3.5a).

Principle co-ordinate analysis (PCoA) using a Bray-Curtis distance matrix at the genus level revealed that bacterial and archaeal communities in sediment samples were different from those in the corresponding water sample, with the two groups clustering separately along the first principal co-ordinate axis (Figure 3.6a). PERMANOVA analysis revealed that sample type alone (sediment or water) explained nearly 30 % of the variance in community composition between samples (see Table 3.2).

Table 3.2: Effect of sample type (sediment vs water) on bacteria and archaea communities across all samples by PERMANOVA. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

	Df	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Type	1	29.355	0.29964	118.94	0.001 ***
Residual	278	68.614	0.70036		
Total	279	97.969	1.00000		

The final microbial eukaryote dataset contained 274 taxa across 295 samples (145 sediment samples and 149 water samples). Like the 16S dataset, the majority of taxa and reads (161, 76.4 %) were found in both water and sediment samples. However, unlike the 16S dataset, water samples contained more taxa overall and more unique taxa than sediment samples (75 unique taxa compared to 38 unique sediment taxa). These unique water taxa made up 20.3 % of the total reads. Water samples had a higher alpha diversity than sediment samples for two out of the three compared measurements (Chao, *t test*, *p adj* < 0.001, Shannon, *t test*, *p* = 0.034, see Figure S3.5 b).

Principle co-ordinate analysis (PCoA) using a Bray-Curtis distance matrix at the genus level revealed that, like the prokaryotic communities, microbial eukaryote communities in sediment samples and water samples clustered largely separately (Figure 3.6b). However, PERMANOVA analysis revealed that sample type (sediment or water) explained around 9.5 % of the variance in community composition between samples (see Table 3.3).

Table 3.3: Effect of sample type (sediment vs water) on microbial eukaryote communities across all samples by PERMANOVA. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

	Df	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Type	1	11.70	0.09465	30.528	0.001 ***
Residual	292	111.91	0.90535		
Total	293	123.60	1.00000		

The multicellular eukaryote dataset was composed of 125 taxa across 292 samples (147 water samples and 145 sediment samples). As with the microbial datasets, the majority of taxa and reads were shared between both types of samples (73, 89.9 %). Sediment samples contained a higher proportion of unique taxa and total reads (35, 8.2 %) than water samples (17, 1.9 %). There were no significant differences in alpha diversity between the two sample types, for all three measures (Chao1, Shannon and Simpson alpha diversity, multiple *t tests* with fdr adjustment).

Principle co-ordinate analysis (PCoA) using a Bray-Curtis distance matrix at the genus and species level revealed that multicellular eukaryote communities and green plant and green algae (Viridiplantae) communities in sediment samples and water samples largely overlapped (Figure 3.6c and 6d). PERMANOVA analyses show that for these larger organisms, sample type is not a dominant factor in community composition, explaining only 3 % of total variance for multicellular eukaryotes (Table 3.4) and 2 % of total variance for green plants and green algae (Table 3.5), although this difference is significant in both cases.

Table 3.4: Effect of sample type (sediment vs water) on multicellular eukaryote communities across all samples by PERMANOVA. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

	Df	Sum of Sqs	R^2	F Model	P(> F)
Type	1	4.164	0.03235	9.796	0.001 ***
Residual	293	124.536	0.96765		
Total	294	128.700	1.00000		

Table 3.5: Effect of sample type (sediment vs water) on green plant and green algae communities across all samples by PERMANOVA. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

	Df	Sum of Sqs	R^2	F Model	P(> F)
Type	1	2.93	0.01925	6.3015	0.001 ***
Residual	321	149.25	0.98075		
Total	322	152.18	1.00000		

The Viridiplantae dataset was composed of 222 taxa spread across 323 samples (180 sediment samples and 143 water samples). Again, most taxa and reads (141, 90.7 %) were found in both types of samples. Water and sediment samples had roughly equal numbers of unique taxa and proportions of overall reads (water: 44, 4 %, sediment: 37, 5.8 %). There were no significant differences in alpha diversity between sediment and water samples (Shannon, Chao1 or Simpson alpha diversity, multiple *t tests* with fdr adjustment).

Season of sampling influences the community composition of water samples more than sediment samples

Figure 3.7 (below): Principal co-ordinate analysis (PCoA) plots of community dissimilarity using a Bray-Curtis distance matrix. Plots in the left column are from water samples, and the right column sediment samples. From top to bottom, the rows contain plots of (i) bacteria and archaea (ii) microbial eukaryotes (iii) multicellular eukaryotes (iv) green plants and green algae. Colours indicate sample month and shapes indicate pond type. Ellipses are shown at the 95 % confidence level.

PCo2 [13.9%]
PCo1 [17.7%]
PCo2 [9.5%]
PCo1 [12.5%]
PCo2 [9.3%]
PCo1 [16.1%]
PCo2 [12.7%]
PCo1 [17%]
PCo2 [10%]
PCo1 [13.8%]
PCo2 [9.5%]
PCo1 [15%]
PCo2 [7.5%]
PCo1 [9.3%]
PCo2 [5.4%]
PCo1 [6.7%]
Pond type EP GP MP SP
Month Ja M Ju S N

For all four different communities, beta diversity was calculated for water and sediment samples separately using a Bray-Curtis index of dissimilarity at the genus level (apart from the Viridiplantae dataset, which was calculated at the species level). A principal co-ordinates analysis was then run and plotted for each (see Figure 3.7). PERMANOVA was run on water and sediment samples for each of the four different communities to determine the impact of sample month and pond type on community composition.

In both sediment and water samples, and across all four different communities, sample month and pond type both had a significant effect on community composition (multiple PERMANOVAs, $p < 0.001$ in all cases, see Tables 6-9). However, in water samples, sample month was a greater influencing factor on community composition, explaining between 10.2 % and 20.2 % of total variation, whereas pond type explained only between 3.3 % and 9.6 %, depending on the community sampled. By contrast, in sediment samples, sample month and pond type had roughly equal effects on community composition, with sample month explaining between 3.4 % and 7.1 % of total variation, and pond type explaining between 4.6 % and 7.1 % of the total variation.

This implies that water communities, as sampled by metabarcoding, experience greater seasonal turnover than sediment communities. Secondly, it implies that, at this scale and for eDNA in water, the influence of the season of sampling on community composition can override the influence of the abiotic conditions on community composition.

Table 3.6: Temporal and spatial effects by PERMANOVA for bacteria and archaea communities across sediment and water samples. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

Water	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	6.1137	0.20199	8.9376	0.001 ***
Pond type	3	2.9049	0.09598	5.6623	0.001 ***
Month:Pond type	9	3.6345	0.12008	2.3615	0.001 ***
Residual	103	17.6141	0.58195		
Total	119	30.2672	1.00000		
Sediment	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	2.730	0.07120	3.1617	0.001 ***
Pond type	3	2.729	0.07117	4.2135	0.001 ***

Month:Pond type	9	2.661	0.06940	1.0272	0.385
Residual	140	30.226	0.78823		
Total	159	38.347	1.00000		

Table 3.7: Temporal and spatial effects by PERMANOVA for microbial eukaryote communities across water and sediment samples. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

Water	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	9.796	0.17095	8.2243	0.001***
Pond type	3	2.936	0.05123	3.2862	0.001***
Month:Pond type	12	6.157	0.10745	1.7230	0.001 ***
Residual	129	38.413	0.67037		
Total	148	57.302	1.00000		
Sediment	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	3.690	0.06758	2.6420	0.001 ***
Pond type	3	2.906	0.05322	2.7739	0.001 ***
Month:Pond type	12	4.360	0.07984	1.0404	0.309
Residual	125	43.647	0.79936		
Total	144	54.603	1.00000		

Table 3.8: Temporal and spatial effects by PERMANOVA on the multicellular eukaryote community across water and sediment samples. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

Water	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	6.506	0.10187	4.1676	0.001***
Pond type	3	2.432	0.03808	2.0770	0.001***
Month:Pond type	12	5.361	0.08394	1.1447	0.051
Residual	127	49.567	0.77611		
Total	146	63.867	1.00000		

Sediment	Df	Sum of Sqs	R^2	F Model	P(> F)
Month	4	3.091	0.05246	2.0193	0.001 ***
Pond type	3	3.585	0.06084	3.1227	0.001 ***
Month:Pond type	12	4.799	0.08143	1.0450	0.284
Residual	124	47.453	0.80527		
Total	143	58.928	1.00000		

Table 3.9: Temporal and spatial effects by PERMANOVA for the green plant and algae community across water and sediment samples. Columns are: sources of variation, degrees of freedom, sums of squares, partial R^2 and p values. One, two or three asterisks are visual representations of p values below 0.05, 0.01 and 0.001 respectively.

Water	Df	Sum of Sqs	R^2	F Model	P(> F)
Month	4	7.385	0.11211	4.5334	0.001***
Pond type	3	2.144	0.03254	1.7547	0.001***
Month:Pond type	11	5.846	0.08874	1.3049	0.001 ***
Residual	124	50.503	0.76661		
Total	142	65.878	1.00000		
Sediment	**Df**	**Sum of Sqs**	**R^2**	**F Model**	**P(> F)**
Month	4	2.856	0.03425	1.6145	0.001 ***
Pond type	3	3.806	0.04565	2.8694	0.001 ***
Month:Pond type	12	5.967	0.07157	1.1246	0.019 *
Residual	160	70.746	0.84853		
Total	179	83.374	1.00000		

Seasonal changes in alpha diversity and taxonomic composition

Microbial communities also exhibited greater differences in alpha diversity between seasons (see Figures S3.6 and S3.7).

For prokaryote communities in sediment, samples from September and November had higher alpha diversity for all three measures (Chao1, Simpson and Shannon) than samples from other months (multiple Wilcoxon tests with fdr p value adjustment, all $p\ adj < 0.05$).

However, for prokaryote communities in water, samples from January consistently had higher alpha diversity than other months (for Chao1, Simpson and Shannon diversity measures, multiple Wilcoxon tests with fdr p value adjustment, all $p\ adj < 0.05$).

In microbial eukaryote communities in sediment, January or March had higher alpha diversity than other months (multiple Wilcoxon tests with fdr p value adjustment, all $p\ adj < 0.05$), whereas for microbial communities in water, January samples were consistently higher in alpha diversity (multiple Wilcoxon tests with fdr p value adjustment, all $p\ adj < 0.001$).

Conversely, multicellular communities did not appear to display differences in alpha diversity across the year (see Figures S3.8 and S3.9). For multicellular eukaryotes in sediment, and green plants and algae in sediment and water, none of the three compared alpha diversity measures showed any significant seasonal differences. However, for multicellular eukaryotes in water samples, September had the highest diversity, and March the lowest (for Chao1 only, multiple Wilcoxon tests with fdr p value adjustment, $p\ adj < 0.05$).

In sediment samples over the year (Figure S3.10), January showed relative peaks of phyla Ciliophora, Bacteroidota, Gastrotricha and Rosales and Poales classes. In March, Annelida, Ochrophyta, Ulotrichales and Asterales were more abundant than in other months, and Ciliophora, Gastrotricha and Poales remained abundant. In July Chlorophyta, and Arthropoda increased in abundance, and Annelida remained abundant. In September and November, Chlorophyta, Arthropoda and Annelida remained abundant. Poales showed a second peak in abundance in September (multiple Kruskal-Wallis tests with fdr p value adjustment, all $p\ adj <0.05$).

In water samples over the year, more prokaryote and eukaryote phyla and green plant orders showed differential read abundance according to sample month (Figure S3.11). In January, Proteobacteria, Cercozoa, Ochrophyta, Basidiomycota, Rotifera, Chlorodendrales and Apiales all showed relative peaks. In March, Ciliophora and Lamiales increased in relative abundance, and Proteobacteria, Rotifera and Cercozoa remained abundant. In July, Proteobacteria was still abundant, and Chlorophyta, Ochrophyta and Charophyceae all increased in abundance. Basidiomycota, Rosales and Sphaeropleales reads were also found in higher numbers. September water samples were characterised by peaks in Actinobacteria, Arthropoda, Cryptophyta and Oedogoniales. Chlorophyta remained high. In November, Chlorophyta, Cryptophyta and Actinobacteria still remained relatively high, whereas Cercozoa and Alismatales increased in read abundance.

Discussion

These four datasets provide wide and deep taxonomic coverage of the communities of these ponds, ranging from bacteria of 0.5 µm in length to trees 50 m tall, and in between, revealing many taxa which are missed by traditional pond surveys, such as rotifers, gastrotrichs and chlorophytes. The high-resolution spatial and temporal sampling of these ponds highlights the fluxes in biodiversity and community composition which can be observed between two ponds barely metres apart, or in the same pond but within a period of 60 days.

eDNA metabarcoding reveals subtle differences in pond taxonomic communities compared to rivers and lakes

The bacteria and archaea community composition conformed taxonomically to what was expected, given previous findings in rivers and lakes (see Introduction), apart from the presence of Desulfobacterota in sediment samples, which has not been previously observed in rivers and lakes. This may be due to recent classification changes creating the phylum Desulfobacterota (Waite et al. 2020). It may also indicate ponds' similarity to freshwater wetlands: Desulfobacterota, which are predominantly sulphate-reducing bacteria, have been found in freshwater wetlands (Prasitwuttisak et al. 2022), and are thought to be responsible for these environments' high rates of sulphate-dependent anaerobic methane oxidation (Cui et al. 2014, Segarra et al. 2015). Ponds and lakes and wetlands are all significant sources of methane, but methane emissions increase on a gradient from natural to artificial or impacted ponds (Rosentreter et al. 2021).

The high abundance of Chlorophyta in the microbial eukaryote community, and the relatively low abundance of Dinoflagellata is different from previous findings from lakes and rivers (see Introduction and Tables S3.1 and S3.2). Interestingly, a recent citizen science programme sequencing eDNA from water from ponds in Bristol, London and Newcastle found that Chlorophyceae was the most abundant microbial eukaryote class by number of species (Natural England 2022). It may be that Chlorophytes are particularly suited to freshwater ponds, due to their shallow depth and high penetration of light, and they have been previously suggested as an indicator of pond quality (Celewicz et al. 2022).

The multicellular eukaryote dataset composition was as broadly as expected from previous surveys of lakes, rivers and ponds (Pearman et al. 2023, Zhang et al. 2021, Wilden et al.

2021). Only one of the invertebrate genera recorded in Pinkhill Meadow ponds using traditional visual identification methods were found using 18S primers (*Musculium*). This is probably due to the low eDNA shedding rates of freshwater macroinvertebrates – low levels of overlap with traditionally identified taxa are found even with specialised primers (see Chapter 2, Leese et al. 2021). However, these 18S primers could pick up organisms difficult or impossible to sample visually, such as microcrustaceans, Rotifera, Gastrotricha, and Nematoda. Interestingly, these primers detected stoneworts (genus *Chara)* in high read abundance. These large algae are traditional clean water pond indicator species (Lambert 2009), suggesting the usefulness of this primer pair for indicating ecological quality.

ITS2 promising for metabarcoding freshwater macrophytes, but with drawbacks

Like the results from the June sampling event, the plant and green alga taxa detected using the ITS2 primers had relatively low overlap with those detected using traditional visual pond monitoring methods (35 species). Whether due to incomplete reference databases, due to low eDNA production and transport from some species, or due to the primers themselves, my method was unable to distinguish between multiple members of the same genus found at the site (e.g., *Juncus, Carex)* and did not detect some abundant species (e.g., *Veronica beccabunga).*

Over a quarter of higher plant taxa were cultivated plants. Some of these may be misidentifications due to incomplete reference databases (e.g., eDNA identified as *Mentha x piperita* (peppermint) may have come from *Mentha aquatica* (water mint)). Others like *Secale cereale* (rye) and *Helianthus annus* (sunflower) belong to species which are grown in nearby arable fields and whose pollen may have been carried in on the wind. The third set, eDNA from species such as *Sesamum indicum* (sesame) or *Prunus armeniensis* (apricot), must either be present due to contamination during the sampling, filtering or extraction stages, or possibly could have made its way through the wastewater treatment system and into the River Thames, and subsequently into the ponds at Pinkhill Meadow.

The large proportion of taxa and reads from terrestrial species shows the potential for eDNA from some distance away to be found in pond water and sediments: *Populus nigra* trees are found at the site, but at 110 m from the nearest sampled pond, yet this species eDNA was present in the samples. 17 of the identified species were non-native, which reflects a mixture of genuine identifications (e.g., *Hesperocyparis arizonica*, Arizona cypress, visually identified at the site) and incomplete reference databases (*Sparganium stoloniferum*, a type of bur-reed native to East Asia, most likely *Sparganium erectum*.)

My research also accentuates the benefits of eDNA metabarcoding with multiple primers on identical samples, allowing more taxa to be detected overall. For instance, the 18S primer pair detected Charophyceae*,* (the stoneworts), large algae and traditional indicator taxa of clean water ponds, whereas this taxon was absent from the ITS2 datasets, even though it sits within the Streptophyta. However, the ITS2 datasets always identified a higher number of Embryophyceae taxa to genus or species level. However, there was also evidence of primer bias obscuring ecological patterns: with the 18S dataset, a greater proportion of Chlorophyta reads were found in sediment samples, whereas with the ITS2 dataset, more Chlorophyta reads were present in water samples.

Despite these drawbacks, the ITS2 primer pair shows promise for metabarcoding macrophytes in freshwater environments, a currently little-explored area of research. Although not distinguishing as many species as visual identification methods, these primers could be used at the genus or family level to contribute to an overall pond quality indicator. The contamination from terrestrial pollen could be avoided by sampling pond sediments, not water. Alternatively, terrestrial eDNA from pond water may be useful in studying overall landscape diversity or health.

Sample season and sample media should be considered carefully when designing eDNA sampling strategies

As expected, communities sampled from pond water and pond sediment were significantly different from each other, and this difference was of a greater magnitude for microbes (around 20 % in prokaryotic microbes and 9 % in eukaryotic microbes) than for multicellular organisms (around 3 % for all multicellular organisms and 2 % for the green plant and algae fraction specifically). These findings are in accordance with Ionescu et al. (2022), who conducted a multi-taxa, seasonal study of 67 ponds using eDNA metabarcoding (16S and 18S primers in sediment and water). The authors found that sample medium (sediment vs water) explained 15 % of the difference in community composition (but didn't report the separate proportion for bacteria or eukaryotes).

It could be argued that these differences are due to primer bias, rather than reflecting an ecological reality. However, I think this is unlikely. Although the 18S primers were developed specifically to target aquatic organisms (Mangot et al. 2013), the 16S primers were developed to amplify both aquatic and non-aquatic taxa equally (Caporaso et al. 2012). Furthermore, if primer bias were the case, we would expect to see a similar proportion of variation explained in both the microeukaryote and multicellular eukaryote dataset.

It is possible that eDNA may be transported and mixed differently in sediment and water, or have different degradation rates in the two media, and this may go some way to explaining the different communities found in these two sample types (Barnes & Turner 2016). However, the differential abundance of some taxa (e.g., greater Rotifera abundance in water samples and Annelida in sediment samples, more Actinobacteria and Cryptophyta in water samples) suggests that eDNA is picking up different community composition due to niche separation.

For microbes, multicellular eukaryotes and plants, and in sediment and water, the month the sample was collected was found to have a significant effect on the resulting community composition. This is in line with previous findings that eDNA detects temporal variation in microbial, metazoan, invertebrate and fungal communities (see Introduction). To my knowledge, this is the first study to find eDNA reflecting seasonal changes in the community composition of macrophytes in a freshwater ecosystem.

The higher alpha diversity found in sampling points earlier in the year (January and March) for microbial organisms, especially in water samples, may be an artefact of greater eDNA preservation at lower temperatures (Barnes & Turner 2016). Ionescu et al. (2022) also report the highest alpha diversity in December and March. Alternatively, the flooding event in January may have increased the overall site richness, as reported in other studies (Chopyk et al. 2020).

Why then did the alpha diversity of multicellular organisms not change significantly month-to-month? Traditional morphological studies have shown that richness of macrophytes and macroinvertebrates is highest in the summer and autumn months (Akasaka & Takamura 2012, Hill et al. 2016). In this study, high rates of eDNA degradation in warmer water may be counteracting the signal of increased alpha diversity in the summer and autumn. However, not all taxa followed the same pattern. Alpha diversity of prokaryote communities in sediment samples peaked in September and November. This could be due to increased temperature: in freshwater lakes, prokaryotic richness decreases with increasing elevation, thought to be related to temperature (Zeng et al. 2016). On the other hand, the increase in alpha diversity at this time may be related to disturbance from nutrient inputs (Beattie et al 2020). Multicellular eukaryotes were richest in water samples in September, as would be expected from traditional macroinvertebrate sampling.

The “pond type” (the 20 ponds on site were divided into one of four pond types depending on location, substrate, water source and vegetation cover) also had a significant effect in all cases. However, whereas in water samples season was the dominant factor, in sediment

samples, season and pond type explained a roughly equal proportion of the variation in community composition. This pattern was found in all four sections of the pond community sampled: prokaryotic microbes, eukaryotic microbes, multicellular eukaryotes and plants and green algae. Therefore, at this small scale and in water samples, the season of sampling obscures finer differences in community composition caused by pond location or abiotic factors. Sediment communities, by contrast, appear to be more stable year-round and experience less turnover. These findings are also in accordance with Ionescu et al. (2022), who found roughly equal effects of land use and seasonality in sediment samples, but a greater impact of seasonality in water samples.

This finding is highly significant for designing eDNA sampling schemes for ponds and other freshwater habitats. Current eDNA pond sampling at scale has involved sampling pond water, not sediments (Natural England 2020, 2022, Harper et al. 2019). To obtain a more complete picture of the pond ecological community and ecological health, I recommend taking both water and sediment samples, and sampling over multiple time points. However, if constraints prohibit this, I recommend taking only sediment samples, especially if sampling several ponds is spread over several months, as this research has shown sediment communities to be more stable throughout the year. Furthermore, only sampling sediments eliminates the costly and time-consuming filtration step.

Conclusion

This study has shown that eDNA metabarcoding is spatially and temporally sensitive on a fine scale, reflecting changes in community composition between the different ponds in this small 5 ha site, and between different times of the year. The clear differences in community composition between sediment and water samples, and in different seasons point to a need for careful design of eDNA sampling strategies to ensure studies into the ecological health of ponds or other freshwaters are comparing like for like.

I found that eDNA metabarcoding of pond water and sediments gives a broader picture of the ecological communities of ponds than traditional methods, detecting organisms rarely sampled in ponds before such as bacteria, algae and zooplankton, and characterising the ecological communities present in pond sediments which were hitherto unknown. These new detections raise the possibility of new pond quality indicator species, such as Chlorophyta

and microcrustaceans. Further research is needed into microbial, microfaunal and algal communities of freshwater ponds to discern whether my findings are site-specific or more universal. Macrophyte communities were successfully detected in ponds using eDNA metabarcoding with a universal ITS2 primer, and displayed distinct community differences by season and pond abiotic conditions. However, the influence of the surrounding terrestrial landscape, particularly in pond water samples, was evident.

Chapter 4: Evaluating environmental and spatial drivers of community composition in lowland ponds: insights from eDNA metabarcoding.

Abstract

This chapter reports results from an investigation into using whole-community, multiple-primer eDNA metabarcoding as bioindicators of freshwater pond ecological quality. Water and sediment samples were collected from 31 lowland ponds of similar geology but situated in varying landscapes (suburban, arable, grassland and woodland) in Southern England in June 2022. eDNA was extracted from these samples and amplified with three different primers to characterise the prokayotic (16S rRNA), eukaryotic (18S rRNA) and green plant and algal (ITS2 gene region) communities of these ponds. Pond physio-chemical and physical variables were measured, and a Habitat Suitability Index (HSI) for Great Crested Newts was calculated for each pond. Spatial autocorrelation was calculated for biological communities of prokaryotes, eukaryotes and green plants, and for environmental variables.

Prokaryotic communities emerged as the most promising bioindicators of pond quality, with community composition in both pond water and sediment differing significantly between ponds of different HSI Levels, with pond water physio-chemistry and with changes in predominant land cover within a 1 km radius. By contrast, eukaryotic communities only showed differences in structure according to surrounding land cover. Green plant and algal communities as characterised by eDNA metabarcoding did not respond to HSI Level, physio-chemistry or land cover. Alpha diversity of green plants and algae was higher in ponds outside of nature reserves, which I posit is due to the interaction of shade, water temperature

and eDNA degradation rates. Green plant and algal communities were spatially autocorrelated, indicating dispersal limitation.

These results have implications for molecular biomonitoring of ponds and other freshwater environments, pointing towards more widespread use of prokaryotes as indicators of ecosystem health and highlighting the effects of pond physio-chemistry on eDNA degradation and alpha diversity estimates.

Introduction

Ponds (lentic freshwater habitats <5 ha in area and <5 m deep, Richardson et al. 2022) are vitally important for freshwater biodiversity and provide many diverse ecosystem services (Biggs et al. 2005, Céréghino et al. 2007, Céréghino et al. 2014). Whilst small individually, collectively ponds are biodiverse, and several studies have found that at a landscape level, ponds contain greater gamma diversity than an equivalent area of rivers or lakes (Oertli et al. 2002, Williams et al. 2004, Davies et al. 2008, Bolgovics et al. 2019). Recent evidence suggests that ponds in agricultural landscapes have positive "spillover" effects for terrestrial species (farmland birds: Lewis-Phillips et al. 2019 and 2020; pollinators: Walton et al. 2020) and adding clean water ponds to an agricultural landscape increases plant richness across a catchment, reversing background declines (Williams et al. 2019).

In the UK ponds have undergone a historic decline in the last century, approximately halving in numbers and density (Biggs et al. 2005, Smith et al. 2022). Aquatic plants, agricultural birds and other wildlife are currently undergoing catastrophic population declines across the UK, linked to changes in land use and management (State of Nature 2023, Williams 2018, 2010). The creation of ponds in agricultural landscapes has been part of agri-environment schemes across Europe and other western countries for the past 35 years as an attempt to reverse these biodiversity declines (DeClerck et al. 2006, Reyne et al. 2021).

Currently, there is no systematic monitoring of pond habitats across the UK by environmental regulators, save the Habitat Suitability Index ("HSI") for Great Crested Newts ("GCNs," *Triturus cristatus*): a protected species and a traditional indicator of a high quality, biodiverse pond (Sewell & Griffiths 2009, Harper et al. 2019). Furthermore, there is limited understanding of what forces structure the biological communities of lowland freshwater ponds: are they structured by environmental factors (either of the pond itself and/or the surrounding landscape), or by dispersal limitation, or both, or neither? This is question

demands further investigation to discern whether freshwater ponds are acting as they are intended in agri-environment schemes: as refugia for biodiversity in highly altered landscapes (Biggs et al. 2017). Reducing nutrient inputs, increasing buffer zones and improving management strategies may all be beneficial for pond biodiversity, but if natural colonisation is limited, for example, due to lack of landscape connectivity, species declines may still be the result (Fahrig 2013, Horváth et al. 2019).

The drivers of macrophyte community composition in lowland ponds have been relatively well studied (Joye et al. 2006, Akasaka et al. 2012, Alahuhta et al. 2014, García-Giron et al. 2019), but by using visual-based identification methods, not eDNA metabarcoding. Conversely, the community assembly of prokaryotic and eukaryotic microbes, zooplankton and phytoplankton has been studied in many different environments using eDNA metabarcoding (e.g. Pearman et al. 2023, Cruaurd et al. 2019 & 2020, Gweon et al. 2019, and see Chapters 1, 2 and 3), but rarely in lowland freshwater ponds and other small wetland habitats.

The relative influence of environment and spatial factors is likely to vary for different organisms (for example, prokaryotes vs eukaryotes, algae vs macrophytes, and communities in pond sediments vs pond water), due to differences in generation times, body sizes, dispersal abilities and life histories (Heino et al. 2015). The relative influences will also differ in different landscapes and at different spatial scales (Heino et al. 2015). One common hypothesis which has support is that smaller organisms such as microbes will be more influenced by local environmental conditions than spatial factors ("everything is everywhere but the environment selects", O'Mally 2008) whereas larger organisms such as macrophytes will be more spatially structured (Padial et al. 2014). Similarly, it is thought that mobile species respond to the environment more than non-mobile species which are dispersal limited and so spatially structured (De Bie et al. 2012).

In reality, empirical results often do not follow these theoretical patterns. For macrophyte and green alga communities, there is conflicting evidence on the importance of environmental factors and spatial structuring. For example, Garcia-Giron et al. (2019) concluded that spatial and environmental factors together determined the community assembly of macrophytes in 51 lowland ponds in a 94,000 km ^{2} region of northwestern Spain, with dispersal limitation interfering with species' tracking of local environmental conditions, such as total phosphorus. However, Gallego et al. (2013) found that in 87 lowland ponds in a 90,000 km ^{2} region of Andalusia, pond water chemistry and pond structure and management were more influential in controlling community structure than spatial factors. Contrary to expectations,

phytoplankton communities displayed more spatial structuring than filamentous green algae or macrophyte communities. Sun et al. (2019), studying macrophytes in 1000 lakes in the UK, found that species were strongly spatially structured. However, environmental variables had more influence in upland lakes than lowland lakes, which the authors suggested was due to increased hydrological connectivity.

Other studies have compared the relative importance of different environmental variables on macrophyte community assembly. In a recent study of 88 lowland agricultural ponds in Northern Italy, Bolpagni et al. (2020) found that total macrophyte richness and the richness of priority species increased with pond area, increasing freshwater connectivity, and urban land cover, but found no effect of other landscape variables. Declerck et al. (2006) measured vegetation complexity as a proxy for macrophyte community richness in 126 farmland ponds in Belgium, and found that it was negatively associated with arable cover in the surrounding 200 m, and trampling of the pond margin by livestock.

Communities of bacteria and archaea in small wetlands seem to conform to the expectation of being environmentally sorted at large scales: in subarctic thaw ponds, bacteria communities were structured by pH, dissolved oxygen, suspended sediments and dissolved organic carbon (Crevecouer et al. 2015), and were not dispersal limited, being either stochastically assembled or environmentally filtered (Comte et al. 2015). Other studies have found assemblages of bacteria communities changed along an elevation gradient, which is in turn related to dissolved nutrient concentrations (Li et al. 2017, Hayden & Beman 2016). However, a study examining the bacterial community composition within a freshwater montane pond found spatial structuring explained most of the variation, and communities did not track environmental variables (Lear et al. 2014). There were differences in community composition observed at >20 m, yet the functional composition did not alter significantly.

Very few studies exist looking at bacteria communities in lowland ponds. A recent study by Ionescu et al. (2022) sequenced bacteria and eukaryote communities in 67 kettle hole ponds in lowland Germany using 16S rRNA gene and 18S rRNA gene eDNA metabarcoding. They found that 5 % of the variation in bacteria community structure in water was related to land use (grassland, arable or forest), but land use was not significant for communities in sediment. Physical and chemical water properties explained 23 % of the variation in communities overall in a constrained ordination, but unfortunately this was not separated for bacteria vs eukaryotes, nor sediment vs water samples.

The same study found that 14 % of the variation in eukaryote communities in sediment could be explained by land use type. No spatial autocorrelation was found for either bacteria or

eukaryote taxa. Looking at unicellular eukaryotes (protists) only, environmental processes were dominant in structuring communities across various scales, from lowland ponds in a small (9.5 km span) regional park (Simon et al. 2015), to montane pools across a 1800 km^2 area (Macingo et al. 2019). Zooplankton communities in a lowland pond network over only 200 ha were found to be environmentally structured (Cottenie et al. 2003), whereas zooplankton community structure in Arctic ponds over 11,475 km^2 was found to be more explained by spatial than environmental factors (Symons et al. 2014). Finally, phytoplankton communities in montane lakes across a 150,000 km^2 area in China were predominantly spatially structured, although environmental factors also played a part (Chang et al. 2021).

eDNA metabarcoding is moving away from research and development and towards becoming a standardised biomonitoring tool with the ability to measure biodiversity more rapidly, for more taxonomic groups, and at greater scales than conventional methods (Deiner et al. 2017, Schenekar 2023). However, many studies have shown low overlap between the species detected by conventional biomonitoring methods and eDNA in aquatic habitats (e.g., Elbrecht et al. 2017, Rivera et al. 2018, and see Chapter 3). Several new promising studies abandon identifying specific bioindicator taxa entirely, instead adopting a "taxonomy-free" approach: linking whole communities sampled via eDNA metabarcoding to known disturbance gradients or biotic indices (reviewed in Corder et al. 2020 and Pawlowski et al. 2018). This method has been effective for bacteria in benthic sediments and the disturbance caused by salmon aquaculture (Stoeck et al. 2018), and diatom 18S rRNA sequences from biofilms in rivers and the benthic diatom index (Apotheloz-Perret-Gentil et al. 2017). The next steps to apply this approach more widely and for different habitats will involve taking widespread biotic indices, benchmarking them against community data and metrics generated by eDNA metabarcoding, and then testing these over larger spatial and temporal scales.

Natural England's Habitat Suitability Index (HSI) is an obvious choice for applying this novel biomonitoring approach to ponds. In this survey, ten indices are calculated for each pond (including physical measurements like area and depth, measurements of water quality, and biotic features such as vegetation cover), and combined to make a composite index which is then categorised as one of five levels, ranging from "poor" to "excellent" habitat for Great Crested Newts (ARG UK 2010). To date, it has been used for over 20 years in thousands of ponds across the UK countryside (Buxton et al. 2021), and records for over 5,800 ponds surveyed between 2017 and 2019 are publicly available (Natural England 2020). Currently,

the HSI records are linked with the collection of pond water, and a qPCR assay applied to determine great crested newt (GCN) presence/absence (Rees et al. 2023).

Recently, HSI levels were correlated with vertebrate species richness and presence/absence determined via eDNA metabarcoding of pond water (Harper et al. 2019). However, the relationship between HSI and the wider components of freshwater pond ecosystems is not understood. This is the first study which examines the relationship between HSI and other taxonomic groups: bacteria, micro-eukaryotes, meiofauna and macrophytes in ponds.

In this study, I sampled water and sediment from 31 lowland ponds across an area of southern lowland Britain of 3,600 km ^{2} in June and July 2022: 17 located within nature reserves, and 14 located outside of nature reserves, within agricultural landscapes. Each sample had the environmental DNA extracted from it, which was then amplified using three separate primer pairs to target three different sections of the pond's biological community: 16S rRNA gene (bacteria and archaea), 18S rRNA gene (eukaryotes) and ITS2 gene region (Viridiplantae - green plants and algae). Water chemistry and physical and structural variables were measured for each pond, and the surrounding land cover at a 100 m and 1 km radius from the pond was extracted from land cover maps. The number of macrophyte species and the percentage coverage of different vegetation types was also estimated in the field.

These environmental variables were used to calculate the Habitat Suitability Index (HSI) for Great Crested Newts (*Triturus cristatus),* and PERMANOVAs and random forest indicator taxa analyses were used to identify whether bacteria, eukaryote and green plant community assemblages and taxa were indicative of different HSI levels. Environmental variables were used separately in variation partitioning analyses to investigate the relative effects of pond water chemistry, pond physical variables and surrounding land cover on the assemblages of bacteria, eukaryotes and green plants. Spatial autocorrelation for environmental variables and bacteria, eukaryote and green plant communities was also calculated, to observe whether there were distance-decay relationships.

The following hypotheses were formulated:

1. Bacteria communities in the ponds will be predominantly environmentally filtered, structured by pond water chemistry variables. Plant and green algae communities will be mainly spatially structured due to dispersal limitation. Eukaryote communities will be structured by a mix of environmental and spatial effects.

2. Pond taxonomic diversity will reflect the quality of the surrounding habitat.
3. There will be differences in community composition of bacteria, eukaryotes and green plants between ponds of different HSI levels.

Methods

Sample ponds

Thirty-one ponds were sampled across Oxfordshire, Buckinghamshire and Northamptonshire in the southern midlands of the UK between 8th June and 27th July 2022. The ponds were chosen to provide a range of landscape types within an area of broadly similar lowland geology and geography. Fourteen ponds outside of nature reserves were selected based on prior PSYM surveys (Predictive System for Multimetrics: an assessment of the ecological quality of a pond based on invertebrate families and plant species present in a pond, Freshwater Habitats Trust 2002) and expert advice (Biggs, J., personal communication, 25th Feb 2022). These ponds are found within areas of predominantly arable or improved grassland land cover and are not managed for biodiversity (see Table 4.1). The other sixteen ponds were located within nature reserves, and managed for biodiversity by wildlife charities. Some of these ponds were in nature reserves in the countryside and surrounded by woodland and neutral grassland, but a large number were found within urban or suburban landscapes (see Table 4.1).

Figures 4.1 and 4.2 below show the locations of the high-impact and low impact ponds. The ponds outside nature reserves were located in two clusters: across a 65.6 km^2 area of Buckinghamshire (b) and a smaller cluster in Northamptonshire (7.4 km^2, a). The ponds inside nature reserves were spread over a larger 418.7 km^2 area of Oxfordshire (c).

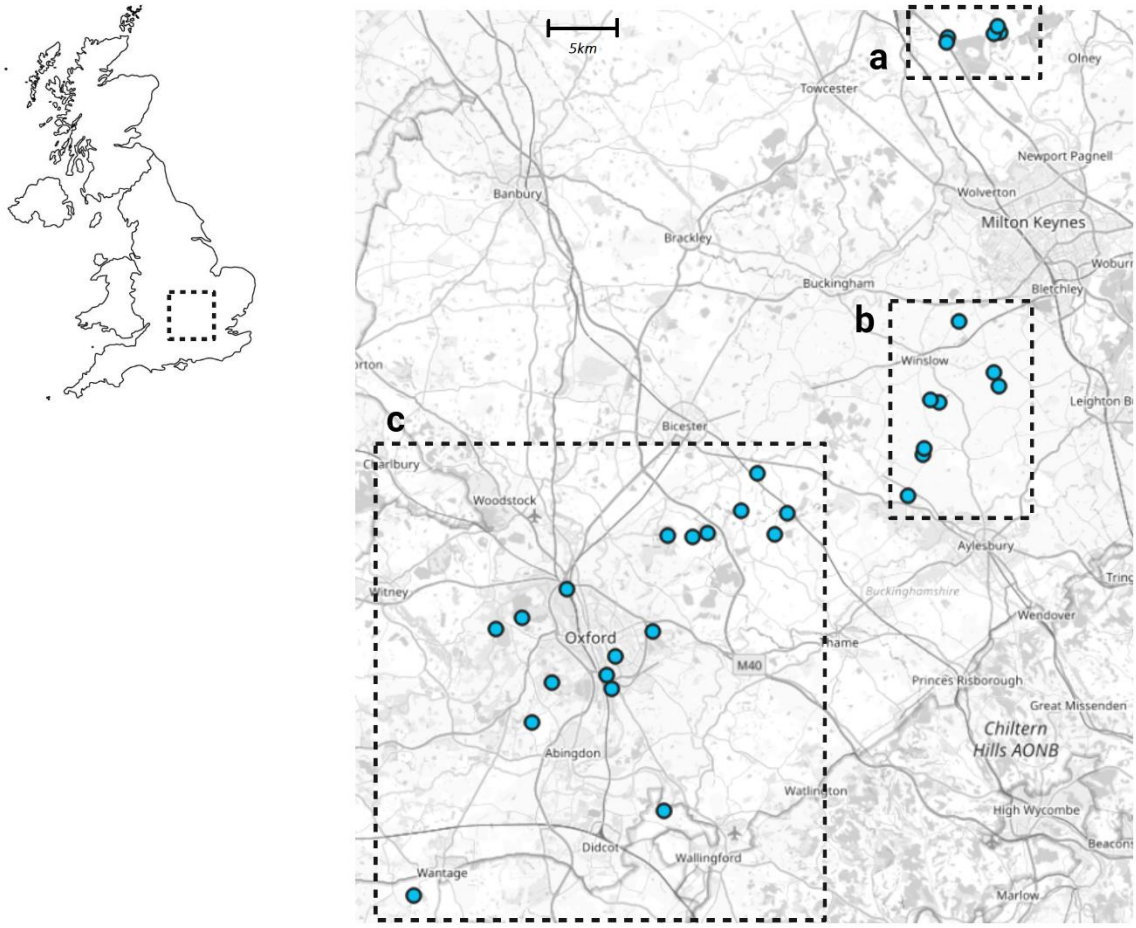


Figure 4.1: The location of the 31 sample ponds within the United Kingdom (left) and within mid-southern England (right). Baselayer: OpenStreetMap, scale 1:1,000,000

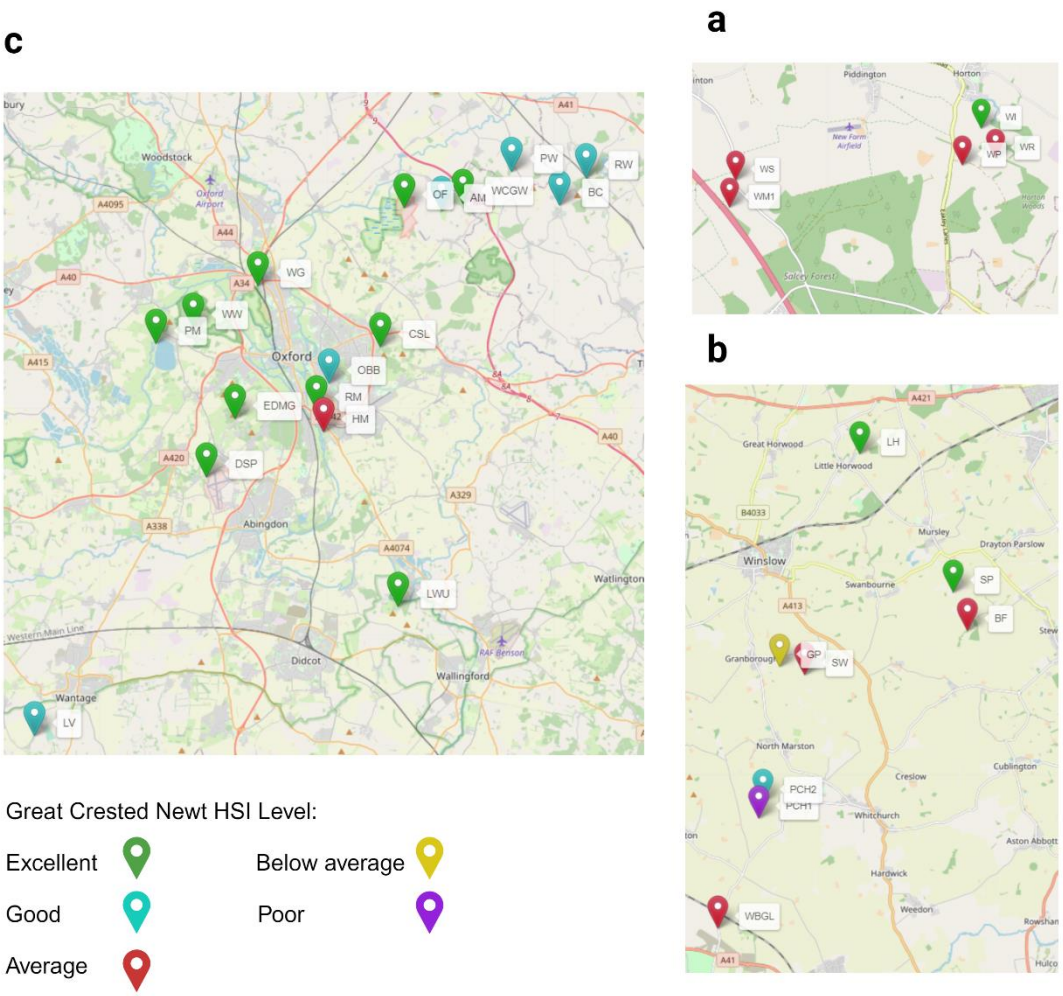


Figure 4.2: Locations and Great Crested Newt Habitat Suitability Index level of 31 sample ponds in a) Northamptonshire, b) Buckinghamshire and c) Oxfordshire. Baselayer: OpenStreetMap, scale 1:500,000 (c), 1:250,000 (b), 1:100,000 (a).

Table 4.1: List of sample ponds. "Nature reserve" column indicates if the pond was located inside a nature (1) or outside (0) of a nature reserve. "Top land cover" columns are the most abundant land cover class within a 100 m or 1 km radius of the centre of the pond by percentage according to the UK Land Cover Map 2021 (Marston et al. 2022). "HSI level" is the Great Crested Newt Habitat Suitability Index level for the pond.

Pond name	Code	Nature Reserve	Top land cover in 100 m radius	Top land cover in 1 km radius	HSI level
ASHAM MEADS	AM	1	Neutral Grassland	Arable	Good
BOAR'S HILL POND	EDMG	1	Suburban	Improved Grassland	Excellent
BOUNDARY BROOK	BB	1	Suburban	Suburban	Good
BRILL COMMON	BC	1	Suburban	Improved Grassland	Good
CS LEWIS NATURE RESERVE	CSL	1	Suburban	Suburban	Excellent
DRY SANDFORD PIT	DSP	1	Neutral Grassland	Improved Grassland	Excellent
FOWL'S PILL	OF	1	Neutral Grassland	Neutral Grassland	Excellent
HEYFORD MEADOWS	HM	1	Grassland	Suburban	Average
LETCOMBE VALLEY	LV	1	Improved Grassland	Suburban	Good
LITTLE WITTEHAM UPPER	LWU	1	Freshwater	Improved Grassland	Excellent
PIDDINGTON WOOD	PW	1	Improved Grassland	Improved Grassland	Good
PINKHILL MEADOW	PM	1	Freshwater	Freshwater	Excellent
RADBROOK COMMON	WW	1	Woodland	Improved Grassland	Excellent
RIVERMEAD	RM	1	Suburban	Suburban	Excellent
RUSHBEDS WOOD	RW	1	Improved Grassland	Improved Grassland	Good
WHITECROSS GREEN WOOD	WCGW	1	Woodland	Improved Grassland	Excellent
WOLVERCOTE GREEN	WG	1	Suburban	Suburban	Excellent
BLACKLAND FARM	BF	0	Improved Grassland	Arable	Average
COWLEYS POND	CP	0	Arable	Improved Grassland	Excellent
GRANBOROUGH POND	GP	0	Arable	Improved Grassland	Below average

LITTLE HORWOOD	LH	0	Improved Grassland	Arable/Improved Grassland	Excellent
PITCHCOTT HILL 2	PCH2	0	Improved Grassland	Improved Grassland	Good
PITCHCOTT HILL FARM 1	PCH1	0	Improved Grassland	Improved Grassland	Poor
SHREW POND	WS	0	Arable	Arable	Average
STEWKLEY POND	SP	0	Improved Grassland	Improved Grassland	Excellent
SWAN'S WAY	SW	0	Arable	Improved Grassland	Good
WADDESDON BLACKGROVE LANE	WBGL	0	Arable	Arable	Good
WOOTTON ICEHOUSE	WI	0	Arable	Improved Grassland	Excellent
WOOTTON M1	WM1	0	Urban	Arable	Good
WOOTTON PONDWEED	WP	0	Improved Grassland	Arable	Good
WOOTTON ROUGH	WR	0	Improved Grassland	Improved Grassland	Good

Pond sediment and water sample collection

One water and one sediment sample were collected per pond in most cases, except for when the area of water exceeded 1500 m^2, when two or three samples were collected (CSL, WI and LWU). This produced 37 water samples and 37 sediment samples. Multiple samples were collected for larger ponds as previous research has shown spatial heterogeneity in communities as detected by eDNA metabarcoding in larger ponds (see Chapters 2 and 3). Water samples were collected using 1L plastic bottles which had been sterilised via a laboratory acid washer or soaking in 10 % bleach. Nitrile gloves were used to collect water and changed between samples. Bottles were triple-washed in pond water before a sample was taken. Water samples were taken around 10cm from the surface of the pond in all cases (a compromise between the tendency of DNA to sink and avoiding disturbing the sediment).

Water was transported in a cool box with ice packs to laboratory facilities within four hours of collection, where it was stored at 4°C for a maximum of 24h before filtration. To detect contamination, two water field blanks were created, where a 1L plastic bottle was filled with 1L of DI water in the laboratory and this was transported to the field and then back to the laboratory.

Sediment samples were collected from the same location as the water samples using a custom sediment sampler comprising of a tubular aluminium rod with a holder for a 50 ml

centrifuge tube, both of which had been sterilised using an acid wash or 10 % bleach before use. The scoop was triple-washed in pond water between each sample. Sediment samples were taken from the surface of the sediment, within the top ~10cm.

To reduce cross contamination between sample sites, the sediment scooper was wiped with 10 % bleach between uses and left to air dry. Sediment samples were transported to the laboratory within 4 hours of collection and stored at -20'C.

Water chemistry measurements

In the field, an ultrameter II (Myron L Company) was used to measure pH, temperature, Total Dissolved Solids (TDS), Conductivity and Oxidative-Reductive Potential (ORP) at each sample point. Measurements were taken at the same time as the water and sediment samples were collected. The sensor was triple-washed in pond water before the measurements were taken. For each sample point, three consecutive measurements were taken, and the mean calculated.

Water depth at each sample point was measured using a metre rule. Water samples for water chemistry analysis were collected using a wide-mouthed bucket which was washed with pond water before sampling. From this larger sample, a smaller 100 ml sample was taken to quantify total suspended solids (TSS) in the laboratory. Two 60 ml samples were subsampled from the bucket and filtered using a syringe and a 0.45um filter in a Swinnex^{TM} (Fisher Scientific, UK) filter holder, on site. The bucket was then stirred, and a third 60 ml sample was collected and not filtered. The two filtered samples were analysed for 1. total Soluble Reactive Phosphorus (SRP) and 2. chlorophyll, ammonia (NH_4), Total Dissolved Nitrogen (TDN) and Total Organic Carbon (TOC). The unfiltered sample was analysed for Total Phosphorus (TP). The water chemistry analysis was carried out using the methods described in Bowes et al. (2018).

Pond habitat

Other pond characteristics were measured using the pond habitat survey method developed by the Freshwater Habitats Trust. This predominantly visual-based standardised survey method is used across the UK to measure the overall habitat condition of ponds (Freshwater Habitats Trust 2015).

Pond outline and permanence were estimated using visual clues such as the winter water line and vegetation changes. Using this as a basis, the percentage of water remaining, and the percentage of the pond overhung with trees and shrubs was estimated. The percentage

coverage of four classes of vegetation (submerged, floating-leaved, emergent and duckweed) was estimated and recorded. The impacts from fish, waterfowl, grazing and other management were visually judged and given a standardised score, and the water clarity was scored similarly. For full details of the method, are found in the survey booklet Freshwater Habitats Trust 2015.

The macrophyte species within the pond area were identified to species level where possible using Greenhalgh & Ovenden (2007). The number of visible macrophyte species was counted and recorded.

Pond area and surrounding landscape

The altitude and area of each pond along with the distance from the pond to the nearest waterbody, and the number of ponds in the surrounding 1 km radius were all measured using tools in Google Earth (version 9.191.0.) The percentage of each type of land cover surrounding each pond was measured using the following method: the locations of each pond were plotted in QGIS (3.26.2-Buenos Aires), buffered to 100 m and 1 km, and shapefiles generated. Using the R packages sf (ver. 1.0-9, Pebesma, 2018) and terra (ver. 1.6-17, Hijmans 2023), the buffers were overlaid on the UK Land Cover Map 2021 (10 m classified pixels, Marston et. al 2022), and the land cover at 100 m and 1 km radius surrounding the pond was extracted. The different land cover classes encountered were arable, coniferous woodland, deciduous woodland, improved grassland, neutral grassland, calcareous grassland, heather, heather grassland, freshwater, suburban and urban.

Great Crested Newt Habitat Suitability Index (HSI)

The Great Crested Newt Habitat Suitability Index (GCN HSI) is a composite index made up of ten different suitability indices, for which a mean is calculated. Scores close to 0 indicate unsuitable habitat, and a score of 1 is optimal habitat. For full details, see ARG UK (2010). An overall HSI score was calculated for each of the 31 study ponds, and then the pond classed as “Excellent” (>0.8), “Good” (0.7-0.79), “Average” (0.6-0.69), “Below average” (0.5-0.59) or Poor (<0.5). The separate suitability indices are:

- SI1: Geographic location
- SI2: Pond area
- SI3: Pond permanence
- SI4: Water quality
- SI5: Shade

- SI6: Waterfowl
- SI7: Fish
- SI8: Pond count within a 1 km ^{2} radius of the pond
- SI9: Terrestrial habitat quality
- SI10: Macrophyte cover

Filtration and eDNA extraction

1000 ml of water was collected for eDNA filtration at each sample point. All samples were filtered in a clean environment using vacuum pump filters sterilised via acid washing or soaking in 10 % bleach. Each 1000 ml sample was initially pre-filtered with a 12 µm Cellulose-Nitrate filter (Whatman, AE100) and subsequently filtered with a 0.45 µm Cellulose Nitrate filter to reduce the influence of large particles of organic matter and plant material on the analysis. The filter papers were removed from the filters using sterilised tweezers and placed in 5 ml centrifuge tubes, which were stored at -20°C until extraction.

DNA from water samples was extracted from the stored filter papers after defrosting using the standard protocol of the DNeasy PowerWater DNA extraction kit (Qiagen Group). An extraction blank was produced by following the normal extraction protocol but omitting any sample. Concentration of DNA was determined using a Nanodrop spectrophotometer (ref).

DNA from sediment samples was extracted by defrosting the full sediment sample and then subsampling 0.25g, using a clean weighing boat and disinfected tweezers. DNA was then extracted using a DNeasy Powersoil kit (Qiagen Group) following manufacturers' instructions. A DNA extraction blank was produced by following the usual protocol but omitting any sample. Concentration of DNA was determined using a Qubit Fluorometer (Invitrogen). All DNA extraction protocols were carried out using sterile procedures i.e. with gloves, and the use of 99 % ethanol and bleach to sterilise equipment and laboratory benches between any handling of DNA samples.

Amplification and sequencing

All DNA amplification and sequencing steps were performed in a laboratory dedicated to environmental DNA analysis, including the use of a room dedicated to PCR preparation, PCR hoods and UV sterilisation, to reduce the risk of contamination.

Three primer pairs were chosen following comparison with results from other primer pairs in a pond environment (see Chapter 2 and 3 for rationale). These primers amplify different DNA or RNA in order to detect different segments of the pond community (Table 4.2 below).

Table 4.2: Primer pairs used in the Pinkhill Meadow seasonal study

Primer pair name and reference	**Amplified fragment length**	**Gene amplified**	**Community primer optimised to target**
NSF573/EKNSR951 (Mangot et al. 2012)	425bp	18S rRNA gene	Eukaryotes
515F/806R (Caporaso et al. 2011 and Walters et al. 2015)	390bp	16S rRNA gene	Bacteria
ITS2-S2F/ITS4_R (Chen et al. 2010)	300-460bp	ITS2 DNA region	Viridiplantae (vascular plants and green algae)

DNA was amplified using a two-step PCR approach. Each primer pair was first tested on 6 randomly chosen samples (three water and three sediment), along with an extraction blank and PCR blank, to assure positive amplification. Then, all 74 samples, including two field and two extraction blanks were amplified with a modified primer (amplicon primers with Illumina MiSeq sequencing primer and pre-adaptor added). Four PCR blanks were also amplified at this stage (normal PCR reagents, but with molecular grade water added rather than any sample). Then step two PCR was carried out to add on the barcodes (Illumina MiSeq index) and flow-cell adaptors. Steps 1 and 2 were repeated for the two other primer pairs. Therefore, each sample had three two-step PCRs carried out on it, one per primer pair. For PCR conditions, see Table S4.1.

Amplicons were normalised using SequalPrep Normalisation Plate Kit, 96-well (Invitrogen, Carlsbad, CA), gel purified using QIAquick gel extraction kit (Qiagen Group) and quantified using Qubit high sensitivity dsDNA Assay kit (Invitrogen, Carlsbad, CA). The resultant amplicon library was sequenced at a concentration of 9 pM with a 0.675 pM addition of an Illumina generated PhiX control library. Sequencing was performed on an Illumina MiSeq platform using MiSeq Reagent Kit v3 (Illumina Inc., San Diego, USA).

Bioinformatic processing

For the 16S rRNA gene and 18S rRNA gene sequences, raw reads were processed through the DADA2 pipeline ver. 1.8 (Callahan et al., 2016) in R (R Core Team, 2018). Briefly, adapters and primers were initially removed from the raw reads using cutadapt (Martin 2011) Then amplicon reads were trimmed to maintain Q score > 30, at 250 and 200 bases, forward and reverse respectively for 16S reads and at 230 and 200 bases respectively for 18S. The

results filtered with DADA2 default settings, except for the maximum number of Ns (maxN) = 0 and maximum number of expected errors (maxEE) = c(5,5).

Sequences were dereplicated and the DADA2 core sequence variant inference algorithm applied. Processed forward and reverse sequences were merged using the mergePairs function, and a sequence table was constructed from the resultant, merged amplicon sequence variants (ASVs). Chimeric sequences were removed from the ASV table using remove BimeraDenovo with default settings.

To assign taxonomy, the online IDTAXA tool from the R package DECIPHER ver. 2.14 (Murali et al. 2018) was used, using the SILVA database, ver. 138.1 (10th March 2021) for 16S and the PR2 database, version 4.13 (17th March 2021), both provided by the DADA2 package. This function gives a taxonomic classification to genus level and provides confidence scores for each level of classification. The confidence threshold for both types of sequence was set at 60 %.

For 16S rRNA gene ASVs, using the R package 'phyloseq' (version 1.42.0, McMurdie & Holmes 2013) samples of less than 4,000 reads were pruned (n = 4), leaving 70 samples. Ultra-rare taxa (less than 3 reads and only in 1 sample) were removed. Samples were rarefied to the median sampling depth (61,159 reads). Removing low-read samples and taxa aids in the removal of false positive detections (Shirazi et al. 2021, García-Machado et al. 2023). Rarefaction is a widely used and statistically valid way to normalize sample size, and so control for the uneven sequencing effort between samples (Weiss et al. 2017, Schloss 2023).

ASVs with the same taxonomic assignment at genus level were agglomerated, and those with a domain confidence of <100 were excluded from the dataset. ASVs assigned to 'Mitochondria' and 'Chloroplast' were also removed. Only two blank samples remained in the dataset after this process (one extraction blank and one field blank), containing 67 taxa. Five taxa within these blanks were present >1000 reads and were also present in environmental samples (unclassifie*d* Comamonadaceae*, Sphingohabdus, Polynucleobacter,* Xanthobacteraceae*,* and *Acinectobacter),* and so were removed from the final dataset. The two blanks were then removed from the final dataset.

For 18S rRNA gene ASVs, no sample pruning was required as all samples had >4,000 reads. Ultra-rare taxa were removed as before, the samples were rarefied to the median sampling depth (24,127 reads), ASVs were taxonomically agglomerated and those of domain confidence <100 were excluded. Eight blank samples in remained in the dataset

after this process, containing 35 taxa (two field blanks, two extraction blanks, four PCR blanks). Five taxa in these samples were present >1,000 reads and were also present within environmental samples (unclassified Mammalia*, Limnostrombidium,* unclassified Annelida*, Synura, Stentor)* and were removed from the final dataset. The eight blanks were then removed from the final dataset.

For the ITS2 region sequences, a different process was followed. Raw reads were processed and taxonomically assigned using the HONEYPI pipeline implemented in Python 2.7 (Oliver et al. 2021). The HONEYPI pipeline removes adaptors and quality filters the raw reads using TrimGalore v.0.6.4, and then uses the DADA2 pipeline to generate an Amplicon Sequence Variant (ASV) abundance table containing chimera-removed, high-quality error-corrected sequences. For each ASV, conserved regions flanking ITS2 are removed with ITSx v.1.1b; and resulting sequences taxonomically classified using the naive Bayesian classifier against an in-house ITS2 database of 966,676 (25th March 2020). Unless stated otherwise, default parameters were used for the steps listed.

Each of the 5,107 ASVs was taxonomically assigned to species level by the RDP classifier within the HONEYPI pipeline and given a confidence level of assignment between 0 and 1.

Using the R package 'phyloseq' samples of less than 4,000 reads were pruned (n=5) and ultra-rare taxa (less than 3 reads and only in 1 sample) were removed. Samples were rarefied to the median sampling depth (23,578 reads). Taxa assigned as "NA", "Fungi" and "Metazoa" at Kingdom level were filtered from the dataset, as were any taxa assigned with confidence <0.97. Taxa were not agglomerated. Three blank samples remained after this process, with 17 taxa (two field blanks and one PCR blank). Of these taxa, only S*alix* was present >1000 reads in blank samples and also present in environmental samples, so this was removed from the final dataset.

All higher plant (Embryophyta) taxa in the pITS dataset (n = 46) were assigned a 'wetland status' function according to the species list used in the National Pond Survey (Biggs et al. 1998). Floating-leaved and submerged species were combined into an "aquatic" category, whereas emergent species and trees or shrubs were combined into a "wetland" category. Plants which did not appear on the NPS list were classified as "terrestrial".

Statistical methods

All statistical analyses were carried out in R, version 4.2.0 (R Core Team 2021). To assess differences in abiotic variables between high-impact and low impact ponds, HSI Levels and

surrounding landscape types, t tests, PCAs and Kruskal-Wallis tests were carried out using functions in base R. Abiotic variable plots were made by using the trans_diff, cal_diff and plot_diff functions in the R package microeco, version 0.11.0 (Liu et al. 2021).

To calculate spatial autocorrelation of pond abiotic characteristics and biotic communities, firstly a distance matrix of Haversine geographic distances between ponds was created using functions in the R package geosphere (ver 1.5-18, Hijmans et al. 2022). Bray-Curtis distance matrices for chemistry variables (12), physical variables (10) land cover at 1 km (10) and 100 m (11) radii were created using the vegdist() function in vegan (ver 2.6-4, Okansen et al. 2022). The same function was used to calculate distance matrices based on read numbers of different ASVs for the separate communities studied: bacteria and archaea (16S dataset), micro-eukaryotes (18S dataset) and green plants and algae (ITS2 dataset).

Mantel tests using Spearman's rho statistic and 9,999 permutations were performed to compare the distance and dissimilarity between pond geographic distances and pond chemistry, physical features, landscape features and prokaryotic, eukaryotic and green plant communities.

Almost all other statistical analyses were carried out in the R package microeco, version 0.11.0 (Liu et al. 2021). Venn diagrams were drawn using the function plot_venn in the trans_venn class. Relative abundance of different taxa was calculated, and bar graphs plotted, using functions within the trans_abund class. Taxa abundance between nature reserve/non-nature reserve ponds, HSI Levels and landscape types were compared via t.test and Kruskal-Wallis test using functions within the trans_diff class. Indicator taxa for nature reserve/non-nature reserve ponds, different HSI levels and dominant land cover at 100 m and 1 km radius were calculated via Random Forests analyses with 1,000 trees and all default settings, applied within the trans_diff class (An et al. 2019). Three measures of alpha diversity (Simpson, Shannon and Chao) were calculated and compared via t.test or Wilcoxon test within the trans_alpha class.

To examine beta diversity, functions within the trans_beta class were used. Beta diversity was calculated using a Bray-Curtis index at the lowest available taxonomic level (genus for bacteria/archaea and eukaryotes, species for green plants and algae), and a Principal Co-ordinates Analysis (PCoA) was performed using the cal_ordination function. A permutational multivariate analysis of variance (PERMANOVA) was performed using the cal_manova function, which uses the adonis2 function in the vegan package, version 2.6-4 (Okansen et al. 2022). Initially, this was performed on all samples using only sample type (sediment vs water) as a factor to examine the effect of sample type on community composition. Then, to

examine the effect of protected status, HSI level, and the dominant land cover at 100 m and 1 km on community composition, sediment and water samples were separated and PERMANOVAs performed. These analyses were marginal: they tested the marginal effect of each term only after accounting for all other effects in the model. This method was chosen as the number of ponds in each class was unbalanced.

To examine environmental drivers of community assembly, firstly a Pearson's correlation matrix of all environmental variables was generated and all strongly colinear variables ($r>0.6$) were removed, leaving 26 variables (nine water chemistry variables, five pond physical variables and twelve landscape variables). Variation partitioning of the Bray-Curtis dissimilarity matrix using the varpart() function in vegan was used to determine the unique and combined fractions of variation in the community explained by the water chemistry, the surrounding landscape and the physical features of the pond habitat. The significance of the six testable partitions were calculated with partial dbRDAs using the capscale() and anova.cca() functions in vegan. The variation partitioning analysis was carried out six times separately: on bacteria + archaea, eukaryotes and green plants + green algae in water and sediment samples respectively.

The variables which were significant in the variation partitioning analyses were then used as constrained variables analyses of community composition. Distance-based redundancy analysis (dbRDA) on the Bray-Curtis matrix of community similarity were performed using functions within the trans_env class of microeco. Environmental vectors were fitted to the ordination using the vegan function envfit(), all default parameters. The ordinations were plotted using the ggord package, version 1.17 (Beck 2022). Other plots were made using functions within microeco which are passed to ggplot2, version 3.4.0 (Wickham 2016).

Results

eDNA concentrations, ASV and read numbers

The concentrations of eDNA in water samples ranged from 2.00 µg/µl to 183.3 µg/µl (M = 35.60 µg/µl, SD = 35.88 µg/µl). DNA was stored in Elution buffer in 1.5 ml microcentrifuge tubes at 20°C. The total volume of extracted DNA ranged from 70 µl – 100 µl. In sediment samples, the range was 13.07 µg/µl to 232.4 µg/µl (M = 56.95 µg/µl, SD = 43.19 µg/µl). DNA

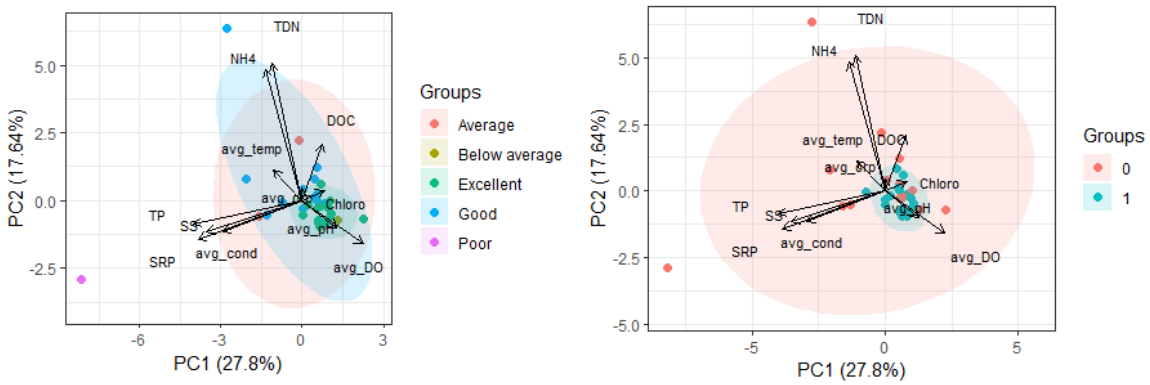
was stored in Elution buffer in 1.5 ml microcentrifuge tubes at –20'C. The total volume of extracted DNA ranged from 80 µl – 85 µl.

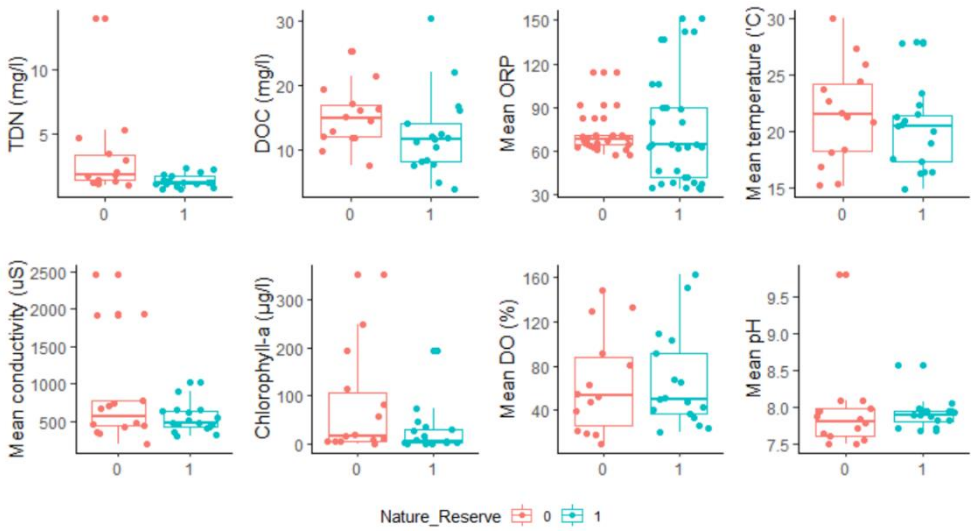
The processing with the DADA2 pipeline (quality trimming, de-replication, chimera removal) produced the following raw datasets: for 16S rRNA sequences, 34,737 ASVs, and the total number of reads was 4,524,825; for 18S rRNA sequences, 2,359,412 merged sequences were recorded in 20,587 ASVs; and for plant ITS2 (pITS) sequences 1,975,561 raw reads and 5,107 ASVs.

After rarefaction, removal of low-read samples and ultra-rare taxa, removal of taxa with low confidence of taxonomic assignment and taxonomic agglomeration (pooling ASVs with the same taxonomic assignment at the genus level), the final 16S dataset contained 979 taxa across 68 samples (34 water samples, and 34 sediment samples), and the final 18S dataset contained 352 taxa across 70 samples (35 water and 35 sediment samples.) The final pITS dataset was not taxonomically agglomerated, and contained 361 taxa across 70 samples (35 sediment and 35 water).

Chemical, physical and landscape characteristics of the ponds

Figure 4.3 (below): Principal component analysis (PCA) plots showing the multivariate variation among 31 ponds in terms of water chemistry variables. Vectors indicate the direction and strength of each chemistry variable to the overall distribution. Coloured symbols correspond to L; the Habitat Suitability Index for Great Crested Newts category (HSI) and R; whether the pond was in a nature reserve (1) or not (0). The first two principal axes explained 45 % of the variance. Avg_pH= mean pH, avg_temp =mean temperature (°C), avg_cond = mean conductivity (µS/cm), ORP = mean oxidative –reductive potential (mV), TDN = total dissolved nitrogen (mgN/l), NH_4 = total ammonia (mg/l), SRP = soluble reactive phosphorus (ug/l), TP = total phosphorus (µg/l), DOC (mg/l) = dissolved organic carbon, SS = suspended solids (mg/l), avg_DO = mean dissolved oxygen (% saturation), Chloro = chlorophyll-a (ug/l)


TDN
NH4
DOC
avg_temp
avg_
Chloro
TP
SS
avg_pH
SRP
avg_cond
avg_DO
PC2 (17.64%)
PC1 (27.8%)
Groups
Average
Below average
Excellent
Good
Poor
Groups
0
1


TDN (mg/l)
DOC (mg/l)
Mean ORP
Mean temperature (°C)
Mean conductivity (uS)
Chlorophyll-a (µg/l)
Mean DO (%)
Mean pH
Nature_Reserve 0 1

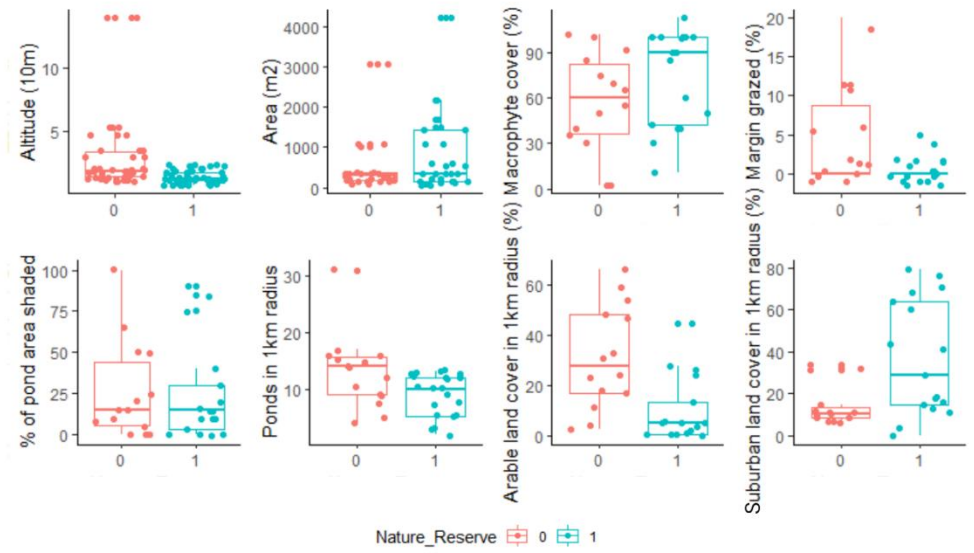

Altitude (10m)
Area (m2)
Macrophyte cover (%)
Margin grazed (%)
% of pond area shaded
Ponds in 1km radius
Arable land cover in 1km radius (%)
Suburban land cover in 1km radius (%)
Nature_Reserve 0 1

Figure 4.4: Box-and-whisker plots of (top two rows) water physio-chemistry variables and (bottom two rows) landscape and physical variables in sample ponds (n=31).

According to PERMANOVA analysis, pond physio-chemistry was different depending on whether the pond was within a nature reserve or not (*df* = 1, *R2* = 0.091, *p* = 0.019) and depending on the HSI category of the pond (*df* = 4, *R2* = 0.286, *p* = 0.026, see Figure 4.3).

Pond physical variables were significantly different depending on the HSI category (*df* = 4, *R2* = 0.334, *p* = 0.007) but not depending on whether the pond was in a nature reserve or not (*df* = 4, *R2* = 0.040, *p* = 0.17). Finally, the landscape in the 1 km surrounding the pond differed between nature reserves and non-nature reserves (*df* = 1, *R2* = 0.089, *p* = 0.029), but not depending on HSI category (*df* = 4, *R2* = 0.134, *p* = 0.37).

No significant differences in individual water chemistry variables were observed between ponds in nature reserves and non-nature reserves when carrying out multiple *t tests* on individual variables with fdr p value adjustment (see Figure 4.4). Ponds outside of nature reserves did have a higher percentage of the margin grazed (*p adj* = 0.032) and were found at a higher altitude (*p adj* = 0.040), whereas ponds in nature reserves were deeper on average (*p adj* = 0.049).

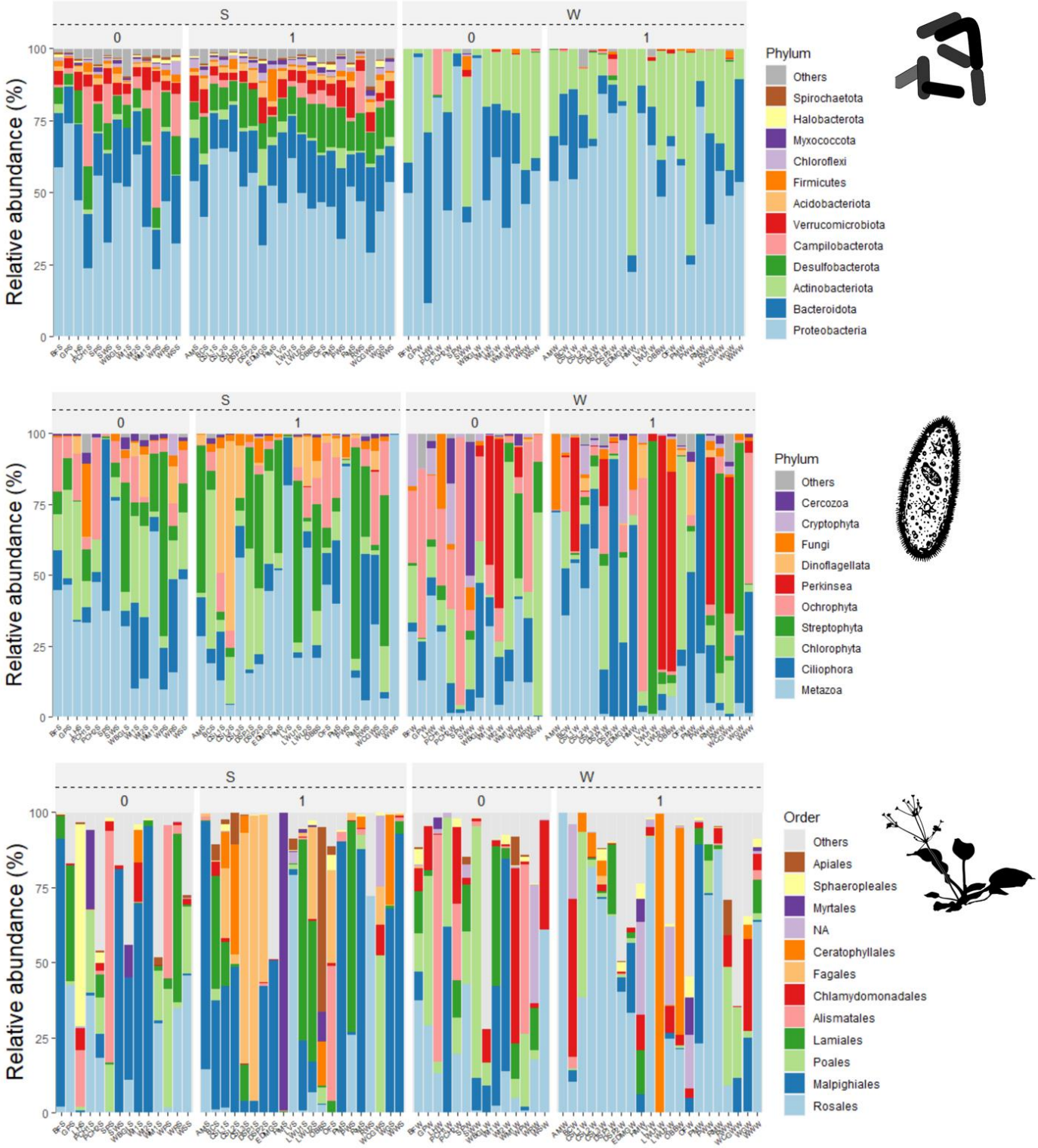
Several landscape variables showed a significant difference between nature reserve and non-nature reserve ponds. Freshwater, suburban and neutral grassland land cover at both scales (100 m and 1 km), heather at 1 km and deciduous wood at 100 m were all higher surrounding nature reserve ponds (all *p adj* < 0.01), whereas arable and improved grassland land cover were more abundant around non-nature reserve ponds (all *p adj* < 0.001). This is unsurprising given landscape features were important factors in choosing the sample ponds. Perhaps more surprising is that non-nature reserve ponds had a significantly higher number of ponds within a 1 km radius (*p adj* = 0.011).

To test the differences in individual variables between different HSI levels, multiple Kruskal-Wallis tests were carried out with fdr p value adjustment. Nine water chemistry variables were found to be different between different HSI levels; however, seven of these were for “poor” or “below average” categories of which there was only one pond of each, and so can be discounted. Total dissolved ammonium (NH_4 mg/l) and total dissolved nitrogen (TDN mg/l) were both found to be significantly higher in ponds classed as “good” (*p adj* = 0.04 for NH_4, and *p adj* = 0.03 for TDN). “Excellent” ponds were found to have a higher area and

average depth (*p adj* < 0.001 for both), and a higher average percentage cover of submerged plants (*p adj* = 0.005). “Excellent” ponds had a higher average distance to the nearest manmade feature (*p adj* = 0.049), coniferous woodland within 100 m (*p adj* = 0.032) and freshwater within 100 m and 1 km (both *p adj* < 0.001). These results are to be expected as different components of the HSI index are based upon the pond area, macrophyte cover and quality of terrestrial habitat.

Taxonomic composition

Figure 4.5 (below): Taxonomic composition of top; bacteria and archaea, middle; eukaryotes and bottom; green plants and green algae communities of 31 lowland freshwater ponds in Oxfordshire, Buckinghamshire and Northants, UK, in June/July 2022. ‘Relative abundance (%)’ = relative read abundance, ‘S’= Sediment samples, ‘W’=Water samples, ‘0‘= pond located outside a nature reserve, ‘1‘= pond located inside a nature reserve. x axis shows sample codes e.g., ‘WCGWS‘ = “Whitecross Green Wood Pond, sediment sample“. For full list of codes, see Table 4.1.



The taxonomic composition of the bacteria and archaea communities of the 31 ponds showed similar patterns to those in previous 16S rRNA gene eDNA metabarcoding studies (see Chapters 2 and 3): other than Bacteroidota, all phyla had a significantly different

percentage read abundance between sediment and water samples (multiple *t tests* with fdr p value adjustment, all *p adj* < 0.05). Water samples were characterised by a higher average proportion of Proteobacteria reads than sediment samples (water: *M* = 59.3 %, *SD* = 20.6 %, sediment: *M* = 48.1 %, *SD* = 12.4 %) and Actinobacteria reads were also higher in water (water: *M* = 24.3 %, *SD* = 17.5 %, sediment: *M* = 2.1 %, *SD* = 2.4 %). Sediment samples had a higher number of average reads from Desulfobacterota (*M* = 10.4 %, *SD* = 3.2 %), Campilobacterota (*M* = 4.9 %, *SD* = 9.0 %), Verrucomicrobiota (*M* = 5.0 %, *SD* = 2.2 %) and others (see Figure 4.5).

The taxonomic composition of eukaryote communities as detected by 18S rRNA gene eDNA metabarcoding showed significant diversity between different individual ponds and between sediment and water samples within each pond (see Figure 4.5). Sequence reads from Metazoa were significantly higher in sediment samples than water samples (sediment: *M* = 36.5 %, *SD* = 24.8 %, water: *M* = 16.1 %, *SD* = 20.1 %, *t test*, *p adj* = 0.008), whereas Perkinsea reads were higher in water samples (water: *M* = 12.5 %, *SD* = 23.6 %, sediment: *M* = 0.3 %, *SD* = 0.05 %, *t test*, *p adj* = 0.04). Specifically, the invertebrate classes Annelida and Gastrotricha, the green algal class Chlorophyta and Xanthophyaceae (yellow-green algae) all had higher average reads in sediment samples compared to water samples, whereas the classes Spirotrichea and Perkinsida had higher average reads in water samples (multiple *t tests* with fdr p value adjustment, *p adj* = 0.045 in all cases).

Looking across sediment and water samples combined, sequence reads assigned to Ochrophyta, and specifically the class Bacillariophyta (diatoms) were more abundant in ponds *outside* of nature reserves (Ochrophyta, non-nature reserve: *M* = 18.8 %, *SD* = 20.7 %, nature reserve: *M* = 9.0 %, *SD* = 14.1 % Wilcoxon rank sum test, *p adj* = 0.024. Bacillariophyta, non-nature reserve: *M* = 6.0 %, *SD* = 7.2 %, nature reserve: *M* = 2.5 %, *SD* = 6.1 %, Wilcoxon rank sum test, *p adj* = 0.011). However, when looking at sediment and water samples separately, no taxa were found to have significantly different abundance at any taxonomic level between ponds inside and outside of nature reserves.

The taxonomic composition of green plant and algae communities across the 31 ponds showed some similarities to eDNA metabarcoding using the ITS2 marker gene in complex ponds in a smaller area (five hectares, see Chapters 2 and 3). Streptophyta sequence reads were more abundant on average in sediment samples (*M* = 92.4 %, *SD* = 14.5 %) compared to water samples (*M* = 79.1 %, *SD* = 24.2 %) whereas Chlorophyta reads were more abundant in water samples (*M* = 20.9 %, *SD* = 24.2 %) than in sediment samples (*M* = 7.6 %, *SD* = 14.5 %, multiple *t tests* with fdr adjustment, *p adj* for both = 0.007).

Overall, in the green plant and algae ITS2 dataset, there were 74 unique Streptophyte species recorded. 42 of these were terrestrial species, 21 were wetland species and 10 were wholly aquatic species. In addition, of the 42 terrestrial species, 10 were trees, and 8 were cultivated. 8 species were non-native. For the full species list including functional annotations, see Table S4.2.

In water samples, eDNA from terrestrial species dominated, making up a mean of 66.5 % (*SD* = 34.9 %) of overall reads, whereas in sediment samples this proportion was 33.9 % (*SD* = 33.8 %) - still a sizeable amount. Wetland species' eDNA made up, a mean of 20.3 % (*SD* = 27.0 %) of total reads in water samples and 54.5 % (*SD* = 35.8 %) of total reads in sediment samples. Finally, reads from aquatic species made up a similar proportion of both types of sample: *M* = 13.2 %, *SD* = 26.5 % in water and *M* = 11.7 %, *SD* = 22.2 % in sediment (see Figure 4.5). eDNA from tree species was found in very high proportions in both types of sample: over a quarter (*M* = 25.7 %, *SD* = 27.8 %) of reads in water samples, and over half (*M* = 57.2 %, *SD* = 36.3 %) of reads in sediment samples. However, eDNA from cultivated species did not make up a significant proportion of reads (*M* = 3.9 %, *SD* = 8.6 % in water samples and *M* = 0.1 %, *SD* = 0.23 % in sediment samples).

Indicator taxa of nature reserves, HSI levels and landscape types

Indicator taxa for nature reserve/non-nature reserve ponds, different HSI levels and most abundant landscape class within 100 m and 1 km radius were identified using a combination of Kruskal-Wallis rank sum tests and the random forest algorithm. No indicator taxa at any taxonomic level from any of the three datasets were found to distinguish between pond communities from nature reserve ponds and non-nature reserve ponds.

135 indicator taxa were found to distinguish between ponds of different HSI Levels: however, these were mostly for poor, below average or average ponds, for which there were only one or three ponds of that classification, and so these taxa were disregarded. Two eukaryote taxa from the 18S dataset were significant indicators of "Excellent" ponds: the Chlorophyte order Sphaeropleales (*p adj* = 0.03) and the Dinoflagellate order Peridiniales (*p adj* = 0.04). Sphaeropleales are vegetatively non-motile colonial or unicellular taxa (Krienitz 2009) and include some of the most common green alga taxa worldwide (Baudelet et al. 2017). Peridiniales (sensu stricto) are a clade of photosynthetic freshwater dinoflagellates (Goméz 2020).

Similarly, most of the 296 indicator taxa found at 100 m or 1 km levels were for rare land classes represented by one or two ponds (e.g., woodland, urban and heathland at 100 m, neutral grassland at 1 km), and were removed from the results. The Gammaproteobacterian genera *Methylobacter* and *Methylovulum* were both found to be indicative of arable land in a 100 m radius (both *p adj* = 0.030). These genera are both methanotrophic (Tveit. al. 2023, Oshkin et al. 2016). Another indicator taxon of 100 m arable cover was an unclassified member of the genus *Nostoc (p adj* = 0.030), a Cyanobacteria genus visible to the naked eye and colloquially known as “witches' butter”. Some species of the *Nostoc* genus are endosymbiotic with hornworts and liverworts (Adams 2002) and some are nitrogen-fixing (e.g., Lindberg 2004).

The species *Ceratophyllum demersum* (Hornwort) was positively associated with ponds with suburban land cover within 100 m (*p adj* < 0.001) and 1 km (*p adj* < 0.001), reflecting the common introduction of this species into managed ponds. *Hedera helix* (common ivy) and *Persicaria amphibia* (amphibious bistort) with ponds with arable land cover within 100 m (*p adj* = 0.012 for both). The colonial green alga *Desmodesmus communis* was weakly associated with arable land cover at 100 m (*p adj* = 0.04).

Alpha diversity

When separating out bacteria and archaea species (16S rRNA gene dataset), sediment samples had significantly higher alpha diversity than water samples for all three measures (Chao, Shannon and Simpson, *t tests* with fdr p value adjustment, all *p adj* < 0.001). Within sediment samples, ponds in nature reserves had significantly higher Chao alpha diversity than ponds outside of nature reserves (*t test*, all p adj = 0.016), however this pattern was not found in water samples. No significant differences in alpha diversity between ponds of different HSI Level or in different landscape types were observed, for either water or sediment samples.

A similar pattern was found in the eukaryote dataset (18S rRNA gene eDNA metabarcoding). Sediment samples had significantly higher alpha diversity than water samples for all three measurements ((Chao, Shannon and Simpson, *t tests* with fdr p value adjustment, all *p adj* < 0.05). This is the opposite of what was found in the seasonal eDNA metabarcoding study in Chapter 3. No significant differences were found, for either water or sediment samples, between ponds inside and outside of nature reserves, of different HSI levels, nor in different landscape types.

For green plant and algae species, no significant differences in alpha diversity were observed between different sample types, but overall samples in ponds *outside* of nature reserves had higher alpha diversity than ponds inside nature reserves, for all three measures (Chao, Shannon and Simpson, *t tests* with fdr p value adjustment, all $p\ adj < 0.05$). When separating out sample types, this pattern held true within sediment samples (for Chao1 and Shannon measures only, *t tests*, both $p\ adj < 0.05$) but not within water samples. The pattern also held for higher plants (Streptophytes, Shannon and Simpson diversity only, *t tests*, $p\ adj$ both < 0.05) but not for chlorophytes. There were no differences in alpha diversity found between different landscape types, at either scale, or HSI Levels.

Another interesting pattern found was ponds with a low number of visible macrophyte species recorded (less than 5) had significantly *higher* Chao alpha diversity of Streptophytes when sampled using eDNA metabarcoding (Kruskal-Wallis test, $p\ adj = 0.016$).

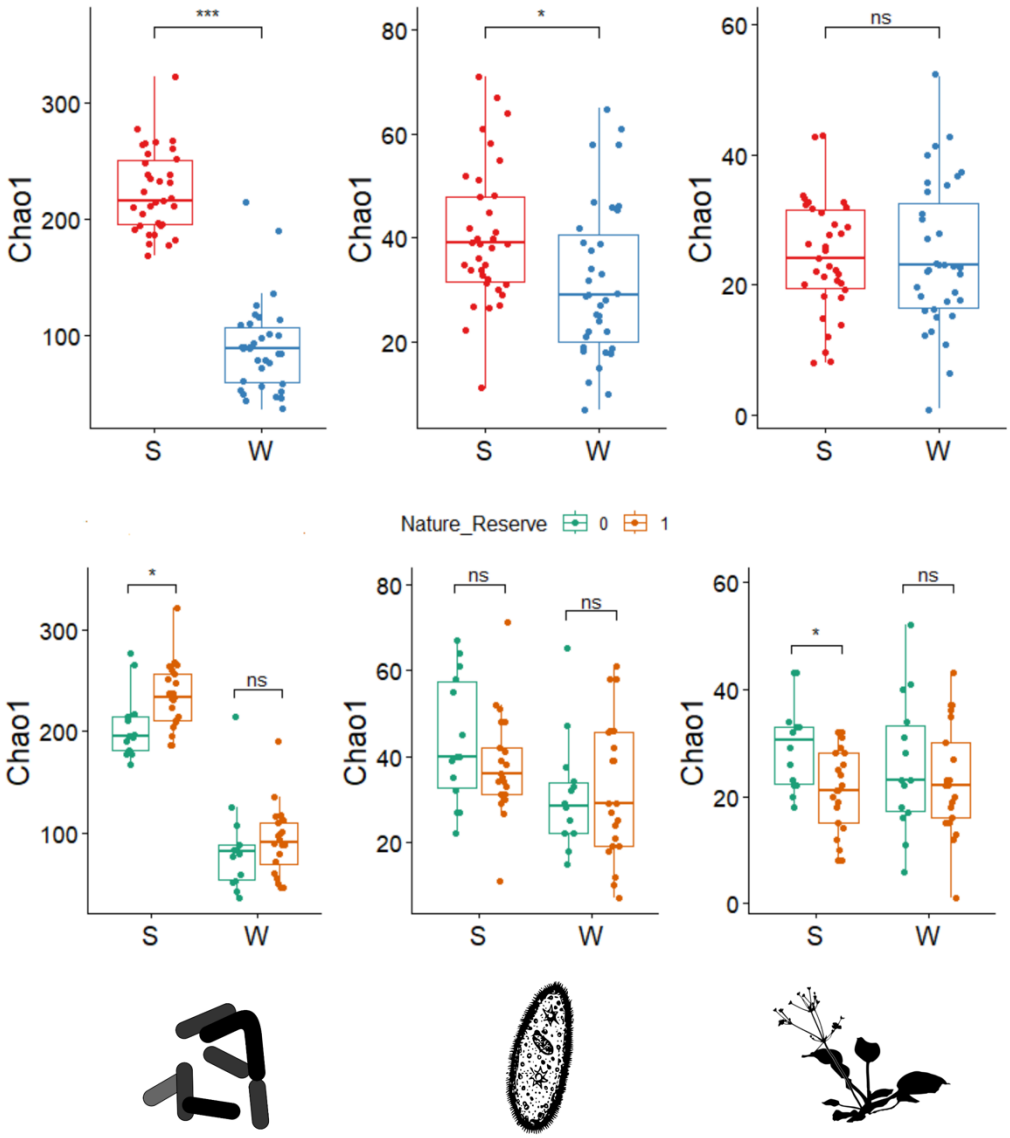


Figure 4.6: Chao1 alpha diversity of L; bacteria and archaea, centre; eukaryotes, and R; green plants and algae in 31 lowland freshwater ponds based on eDNA metabarcoding assays. Top row: comparison between sediment (S) and water (W) samples. Bottom row: comparison between ponds inside (1) and outside (0) nature reserves. One and three asterisks indicate whether the difference was significant at the p<0.05 and p<0.001 level respectively, according to Student's t test. 'ns' = no significant difference.

Spatial autocorrelation of ponds abiotic characteristics and biological communities.

There was a significant spatial autocorrelation of water chemistry variables: ponds closer together had more similar water chemistry than ponds further apart (*Mantel's r statistic* =

0.267, $p = 0.002$). Ponds closer together also had more similar land cover composition within a 100 m radius (*Mantel's r* = 0.312, $p < 0.001$), but not within a 1 km radius (*Mantel's r* = 0.089, $p = 0.092$). However, there was no significant spatial autocorrelation in ponds' physical variables (*Mantel's r* = 0.084, $p = 0.052$).

There was no significant spatial autocorrelation in the bacteria and archaea community amongst ponds (16S dataset, *Mantel's r* = 0.037, $p = 0.30$), nor in the eukaryote community (18S dataset, *Mantel's r* = 0.030, $p = 0.27$). However, there was significant spatial autocorrelation of the plant and green algae communities (ITS2 dataset, *Mantel's r* = 0.018, $p = 0.005$): ponds closer together geographically had more similar ecological communities. This pattern was not observed when separating the samples into sediment and water samples, nor in Streptophyta and Chlorophyta taxa.

Community composition

PERMANOVA analyses on a Bray-Curtis matrix of community dissimilarity were carried out testing for the marginal effect of each term after accounting for all other terms in the model. This method was chosen because some of the categories examined (HSI Level, Landscape type 100 m and 1 km, Macrophyte species number) had a low and uneven number of ponds in each category.

Ponds within nature reserves and outside of nature reserves did not have significantly different community compositions of bacteria and archaea, eukaryotes or green algae and plants, in neither sediment nor water samples.

In bacteria and archaea communities (16S dataset), sample type (sediment vs water) had a large influence on community composition (*R2* = 0.338, $p < 0.001$), a similar proportion to that found in previous studies using the same marker in a smaller area (see Chapters 2 and 3). Sediment and water samples were then separated, and PCoAs and PERMANOVAs carried out separately. In sediment samples, HSI Level (*R2* = 0.142, $p = 0.008$) and the predominant land cover type in a 100 m radius (*R2* = 0.226, $p = 0.02$) had a significant marginal influence on bacteria and archean community composition (see Table 4.3). In water samples, only HSI Level was found to be significant (*R2* = 0.155, $p = 0.03$).

In eukaryote communities (18S dataset) sample type had a much smaller, but still significant, effect on community composition (*R2* = 0.060, $p < 0.001$), but no significant effects of HSI Level or surrounding land cover type were observed in sediment or water samples separately. In plant and green algae communities, sample type again had a small but

significant influence (*R2* = 0.057, *p* < 0.001). No significant factors were found to explain differences in community composition in sediment samples. However, in water samples land cover at 1 km explained 18.4 % of variation in communities (*p* = 0.005), whereas the number of visible macrophyte species explained 10.3 % of total variation (*p* = 0.034).

Table 4.3: R2 values of factors in PERMANOVA analyses for three taxonomic groups in lowland ponds, June 2022. One, two and three asterisks indicate whether the analysis was significant at the p<0.05, p<0.01 and p<0.001 level respectively, and orange shading indicates the factor was significant. "Nature Reserve" = binary factor, if the pond was located in a nature reserve or not. "HSI Level" = habitat suitability index for Great Crested Newts of pond. "Land cover at 1 km or 100 m" = the most dominant land cover in a 100 m or 1 km radius of the pond, using the UK Land Cover Map 2021 (Marston, 2022). "Macrophyte no. class" = estimated number of macrophyte species per pond (0-5, 6-10, 11-15 or 16-20).

	Nature Reserve (df = 1)	**HSI Level (df = 4)**	**Land cover 100 m radius (df = 7)**	**Land cover 1 km radius (df = 5)**	**Macrophyte no class (df = 3)**
Bacteria and archaea (16S)					
Sediment	R2=0.03, p=0.17	R2=0.14, p<0.01**	R2=0.23, p<0.05*	R2=0.14, p=0.32	NA
Water	R2=0.02, p=0.37	R2=0.16, p<0.05*	R2=0.20, p=0.25	R2=0.14, p=0.37	NA
Eukaryotes (18S)					
Sediment	R2=0.02, p=0.85	R2=0.12, p=0.22	R2=0.22, p=0.07	R2=0.14, p=0.39	NA
Water	R2=0.04, p=0.07	R2=0.13, p=0.11	R2=0.19, p=0.62	R2=0.16, p=0.08	NA
Plants and green algae (pITS)					
Sediment	R2=0.02, p=0.67	R2=0.10, p=0.55	R2=0.21, p=0.22	R2=0.16, p=0.11	R2=0.09, p=0.25
Water	R2=0.03, p=0.37	R2=0.11, p=0.30	R2=0.19, p=0.23	R2=0.18, p<0.01**	R2=0.10, p<0.05*

Environmental determinants of community composition

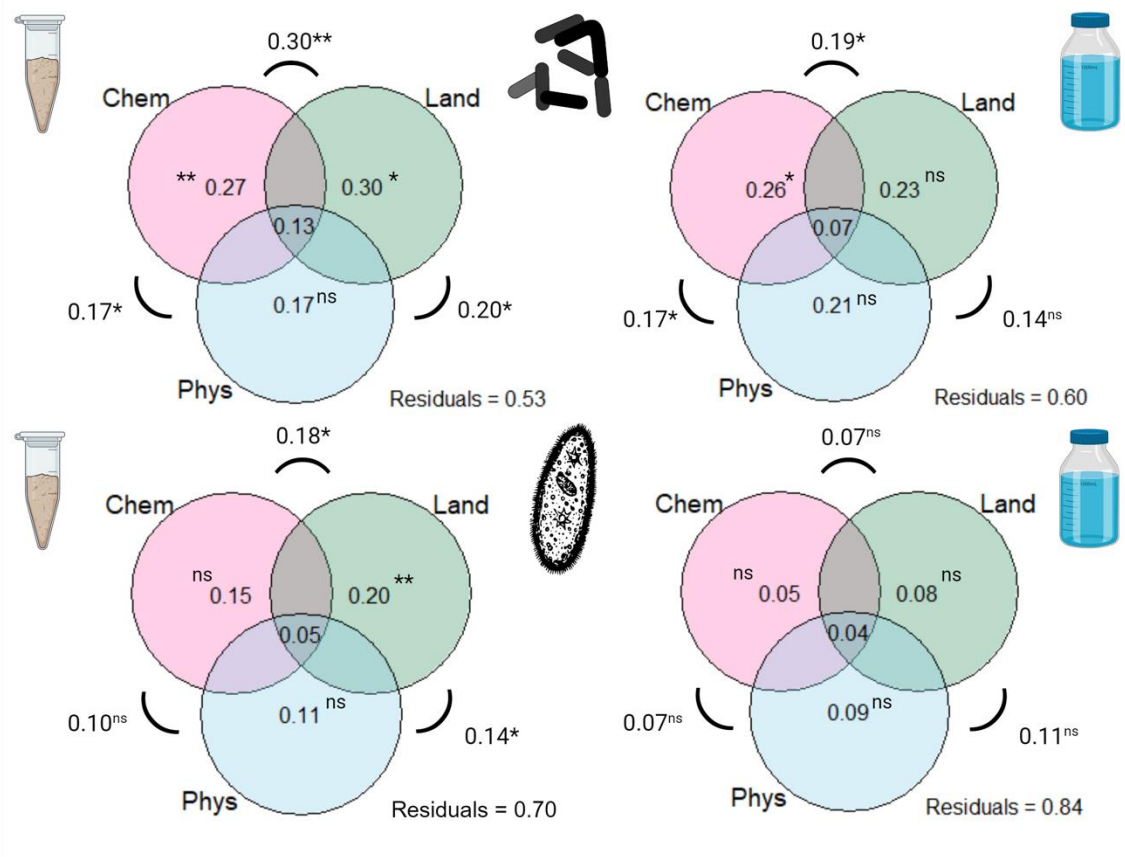


Figure 4.7: The results of variation partitioning analyses for top: bacteria communities, and bottom: eukaryote communities in sediment (L) and water (R) samples. Variation in a Bray-Curtis dissimilarity matrix is partitioned between that explained by chemical (9 measures), landscape (12 measures) and physical (5 measures) variables, and a residual component. Numbers indicate the adjusted R2 value of the variation partitioning analysis. One and two asterisks indicate that the variation explained by the portion was significant when tested by partial dbRDA, at p<0.05 and p<0.01 levels respectively. "ns" indicates that the dbRDA was not significant for that portion. Results for green plant communities are not plotted as all testable portions were not significant.

Variation partitioning was carried out on each of the three communities (bacteria and archaea, eukaryotes, green plants and algae) and for sediment and water samples separately. This was done to investigate the effect of water chemistry, surrounding landscape, or the pond's physical characteristics on the pond's community composition, whilst accounting for the effects of other environmental variables (see Figure 4.7).

In bacteria and archaea communities in pond sediments, water chemistry alone accounted for 27 % of the variation in community structure (F = 3.86, p = 0.004) and surrounding landscape alone accounted for 30 % (F = 1.77, p = 0.02), but the proportion accounted for by physical pond variables was not significant. In bacteria and archaea communities in pond water, water chemistry alone accounted for 26 % of the variation in community structure (F = 2.06, p = 0.018), but landscape and pond physical properties were not significant by themselves.

In eukaryote communities in sediment, only surrounding landscape was found to have a significant influence on community structure, accounting for 20 % of variation in community structure whilst controlling for the other variables (F = 2.66, p = 0.004). In eukaryote communities in water, none of the three components were found to significantly account for the variation in community structure. Similarly, in plant and green algae communities in both water and sediments, the proportions of variation explained by water chemistry, surrounding landscape and pond physical characteristics were not significant.

To investigate the effect of individual variables on community composition, dbRDAs were carried out using the variables (e.g., pond chemistry variables) and communities (e.g., eukaryote taxa in pond sediments) identified as significant in the variation partitioning analysis /partial dbRDAs.

The results of these analyses found that water chemistry variables explained around 9.5 % of the variation in bacteria and archaea communities in water (*R2 adj* = 0.095, see Figure 4.8, top left), the first two axes accounted for 45.6 % of the total variation, and the overall model was significant (F = 1.38, p = 0.02). Conductivity and total phosphorus were significantly correlated with dbRDA2 and dbRDA1 respectively (conductivity p = 0.049, total phosphorus p = 0.017).

Water chemistry variables also explained around 12 % of the variation in bacteria and archaea communities in sediment (*R2 adj* = 0.119, Figure 4.8, top right), the first two axes accounting for, again, 45.6 % of the total variation, and the overall model was significant (F = 1.49, p < 0.001). In this case, conductivity (p = 0.002), mean temperature (p = 0.013), dissolved organic carbon (p = 0.012) and total phosphorus (p < 0.001) were all positively correlated with dbRDA1, whereas dissolved oxygen was negatively correlated with dbRDA1 (p = 0.002).

Landscape variables explained 17 % of the variation in bacteria communities in sediment (*R2 adj* = 0.173, Figure 4.8, bottom left). The first two axes accounted for only 37.6 % of the

total variation, but the overall model was highly significant (F = 1.58, p < 0.001). The percentage of arable and improved grassland land cover within a 1 km radius were positively correlated with dbRDA1 (arable, p = 0.042, improved grassland, p = 0.023), whereas the percentage of heather cover in 1 km was negatively correlated with dbRDA1 (p = 0.016), as was the percentage of suburban land cover (p = 0.002).

Landscape variables explained 10 % of the variation in eukaryote communities in pond sediments (*R2 adj* = 0.097, Figure 4.8, bottom right). The full model was significant (F = 1.31, p < 0.001) although the first two axes only accounted for 31.6 % of the total variation. Number of ponds in a 1 km radius of the study pond was positively correlated with dbRDA1 (p = 0.003), whereas the percentage of arable and urban land cover in the surrounding 1 km were positively correlated with dbRDA2 (arable p = 0.007, urban p = 0.001).

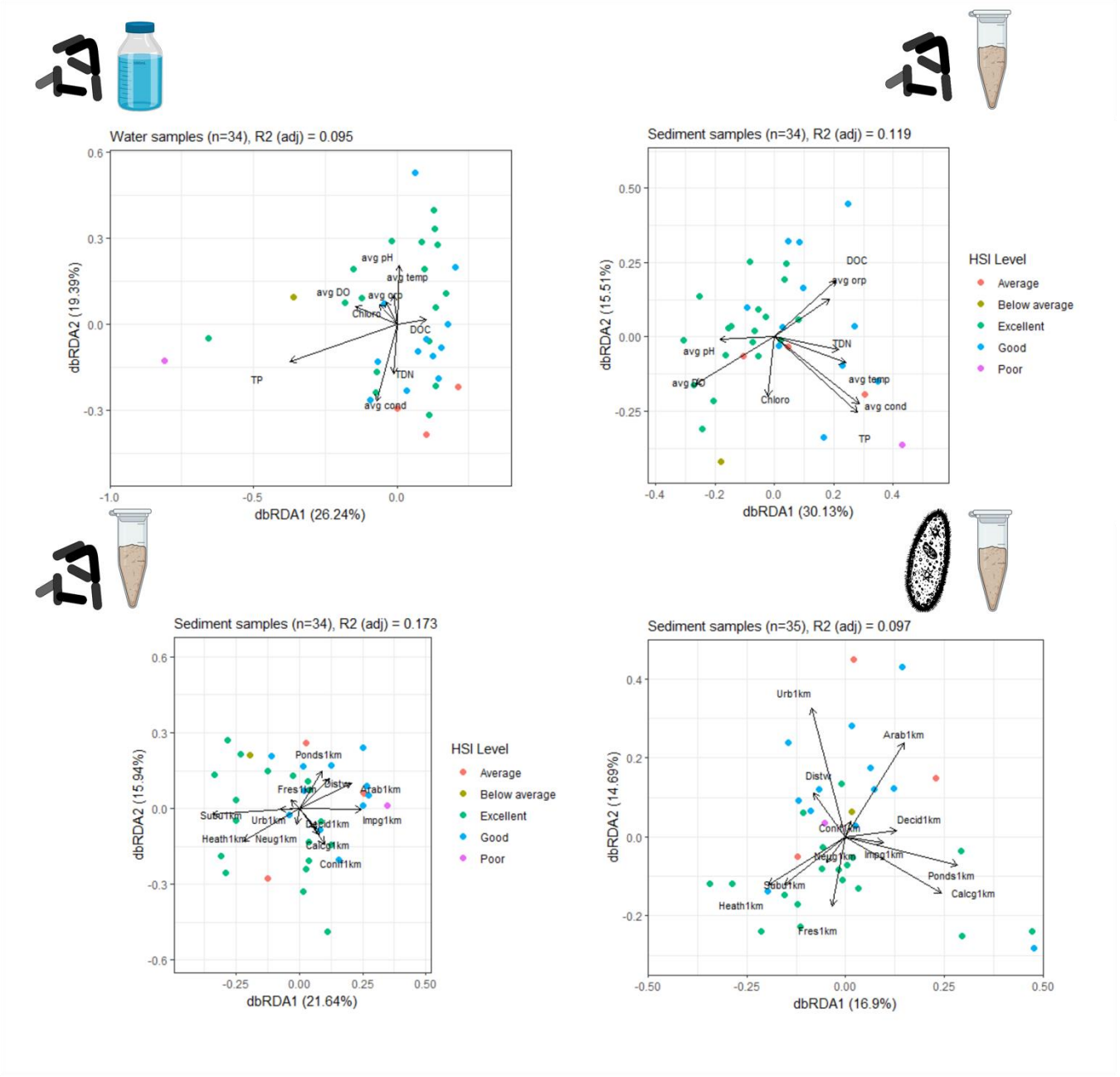


Figure 4.8: Distance-based redundancy analyses (dbRDA) of Bray-Curtis distances between samples from ponds of different HSI Level (colours). Top left: bacteria in pond water and water chemistry variables; top right: bacteria in pond sediments and water chemistry variables; bottom left: bacteria in pond sediments and % land cover within a 1 km radius; bottom right: eukaryotes in pond sediments and land cover within a 1 km radius. The full models of the ordinations were all significant ($p < 0.05$).

Discussion

The findings of this study shed light on the potential of bacteria and archaea communities and taxa as reliable indicators of freshwater pond quality. My investigation revealed discernible variations in bacterial community composition within both sediments and water across different HSI levels, explaining 14-16 % of the variation in communities. Interestingly, ponds located within nature reserves exhibited notably higher prokaryote richness in sediments than those located outside nature reserves.

Unconstrained and constrained ordinations of bacterial community structure yielded consistent results: sediment bacteria communities were structured by pond water chemistry variables and land cover variables, whereas communities in water only responded to water chemistry variables. Specifically, total phosphorus and conductivity (for both water and sediment), and the amount of arable, improved grassland and suburban land cover in the surrounding area (for sediment only) appeared to be important determinants of community structure. Bacteria and archaea communities were not spatially structured.

These findings exhibit support for bacteria and archaea communities in ponds, particularly in sediments, responding to anthropogenic impact, possibly mediated through land cover change and associated changes in water chemistry. Prior research has found evidence of bacteria in freshwater sediments showing signals of wastewater discharge into rivers (Martinez-Santos et al. 2018), dam building (Zhang et al. 2017) and wetland drainage (Wood et al. 2021). In this landscape, Methanogens were indicator species of arable land cover in the immediate vicinity of ponds. Previous studies have demonstrated Methanogens' sensitivity to eutrophication and temperature (Nijman et al. 2021, Yang et al. 2019), underscoring their potential role as biomarkers for environmental changes.

Like previous findings, the alpha diversity of prokaryotes was higher in pond sediments than pond water, and the community composition differed in the two environments, emphasising the importance of sampling encompassing both pond sediments and water to comprehensively capture microbial responses to environmental factors.

Our investigation also explored eukaryotic communities, identifying certain taxa as indicators for both HSI levels (Peridiniales, Sphaeropleales*)* and nature reserve status *(*Ochrophyta, Bacillariophyta*).* The cause of this is difficult to discern as taxa lower than order were not identified. Diatoms may be responding to the higher nutrient levels in ponds outside of nature reserves (Smucker et al. 2020, Kiran et al. 2016), or an increased number of ponds in

the surrounding landscape aiding dispersal, perhaps via waterfowl (Manning et al. 2021, Soininen 2007). However, despite these indicator species, no discernible differences were observed in community compositions according to HSI level or nature reserve status.

The interplay of landscape variables within a 1 km radius was identified as a potential influence on eukaryote community composition in sediments, particularly arable and urban land cover and the number of nearby ponds. However, none of the measured variables (chemical, physical or landscape) were found to have significant structuring effects on eukaryote communities in pond water. Notably, spatial autocorrelation was absent in the studied communities.

My findings agree with Ionescu et al. (2022) in finding a similar proportion of variation in eukaryotic communities in pond sediments explained by land cover. However, no environmental filtering was indicated in eukaryote communities in water samples, contrary to previous findings for protists, phytoplankton and zooplankton (Macingo et al. 2019, Simon et al. 2015), and no spatial structuring was observed in either water or sediment samples, contrary to previous findings (Chang et al. 2021).

It may be that different functional eukaryote communities (e.g., protists, phytoplankton and zooplankton) are structured differently, and it is necessary to analyse these groups separately to tease out the differential patterns. However, other explanations are possible, for instance, homogenising dispersal or homogenising selection (Pearman et al. 2023). Additionally, eukaryote communities in freshwater lakes exhibit a high temporal turnover, and sampling a single timepoint may obscure spatial and environmental effects (Li et al. 2020, Chapter 3). This finding is supported by the fact that in this study, alpha diversity was higher in pond sediments than water, yet in a previous study using the same primer pair over the course of a year (Chapter 3), alpha diversity was higher in pond water.

In contrast, green plants and algae did not produce indicator species for either nature reserves or HSI levels and community composition within this group did not show any significant response to these factors.

Surprisingly, plant richness in sediment was higher *outside* nature reserves, and in ponds with *fewer* visible macrophyte species. This counterintuitive finding may be explained by the interaction between shade, macrophyte abundance and rates of eDNA preservation. On closer examination, ponds with <5 macrophyte species were significantly more overhung by trees than ponds with a higher number of visible macrophyte species (Kruskal-Wallis test, $H(3) = 33.68$, $p < 0.001$, post-hoc Dunn tests all $p < 0.001$). It may be that higher levels of

shade suppress the growth of macrophytes in the pond, but also result in higher rates of eDNA deposition or preservation due to lower temperatures and levels of UV light. However, the pattern was not found in water samples: perhaps the rate of turnover of eDNA in water is higher than in sediment samples, obscuring the effect of shade.

No environmental variables were found to be significant drivers of green plant and algae community structure, either in pond water or pond sediment. However, there was significant spatial autocorrelation in green plant and algae communities. This suggests that the limited dispersal ability of both higher plants and green algae is causing this spatial structuring (Garcia-Giron et al. 2019, Gallego et al. 2013), particularly as the pattern only held true when combining higher plant and green algae communities, and sediment and water samples. However, it may also reflect the high influence of the surrounding terrestrial environment, as land cover within 100 m displayed spatial autocorrelation, and 30 % of the total eDNA reads in sediment samples and 65 % of the total eDNA reads in water samples originated from terrestrial plant species.

It's crucial to acknowledge the potential confounding effects of spatial autocorrelation in water chemistry on the interpretation of responses to HSI levels and nature reserve status. While this spatial autocorrelation was not observed in bacterial or eukaryote community similarity, it poses a significant consideration for future research design.

Taken together, the results of this study show that pond communities detected by eDNA metabarcoding, particularly microbes, can respond to differences in pond quality. Future research should focus on bacterial communities as biomarkers for assessing the ecological health of freshwater lowland ponds. An obvious next step would be to repeat the study over a wider geographical extent, with a greater number of ponds and more replicates in each Habitat Suitability Index (HSI) category. Another potential research direction is developing novel pond health indices by comparing community data derived from eDNA metabarcoding with PSYM scores. Moreover, to gain deeper insights into the spatial and environmental structuring of pond communities, a multifactorial experimental design is recommended: for instance, sampling clean water ponds in areas of low-quality terrestrial habitat and vice versa.

Conclusions

eDNA metabarcoding is a powerful tool for investigating ecological hypotheses in freshwater ponds at large spatial scales, and provides opportunities for future biomarkers in these environments.

In this study, community assemblages of bacteria and archaea emerged as the best indicators of different HSI levels, but not green plant or eukaryote communities.

This suggests that prokaryotes possess the potential to serve as robust biomarkers for delineating ecological quality levels between pond ecosystems. Future research endeavours should aim to test the applicability of this methodology by sampling a greater number of ponds across a broader geographical expanse, incorporating multiple replicates within various Habitat Suitability Index (HSI) categories, to benchmark prokaryote assemblages in excellent, good, average, below average and poor HSI quality ponds. Given the thousands of ponds across the UK with documented HSI levels, this research presents an enticing opportunity.

The bacterial communities observed in both water and sediment exhibited clear structuring according to water chemistry parameters. Moreover, sediment bacterial communities demonstrated sensitivity to land cover within a 1 km radius of the pond. Similarly, eukaryotic communities present in pond sediments displayed structuring influenced by variations in land cover, although no spatial structuring was evident.

This investigation also highlights the limitations of eDNA metabarcoding in accurately portraying ecological patterns concerning macrophyte and green algae communities within freshwater ponds when compared to traditional survey methodologies. Metabarcoding outcomes for plant communities are notably biased by the presence of eDNA from the surrounding terrestrial landscape, and the interplay of shade on eDNA preservation and degradation may lead to conclusions that diverge from those obtained through traditional approaches.

Notably, environmental variables were not identified as significant determinants of structure within green plant communities. However, spatial structuring was evident, suggesting potential constraints due to dispersal limitation or the influence of the surrounding landscape within a 100 m proximity to the ponds, a factor that also displayed spatial autocorrelation.

Consistent with prior research, this study reaffirms the pivotal role of suburban ponds as significant reservoirs of freshwater biodiversity (Hill et al., 2017), while also highlighting the impact of agricultural practices on pond water chemistry and, by extension, biological communities (Sawatzky et al. 2019, Fuentes-Rodrigues et al. 2013, Leibold 1999). Furthermore, there exists preliminary evidence suggesting a positive relationship between freshwater connectivity and heightened eukaryotic diversity.

Chapter 5: Conclusions

This collection of studies represents one of the most thorough applications of eDNA metabarcoding in freshwater ponds to date. These are some of the first studies to use eDNA metabarcoding of the 16S rRNA gene and 18S rRNA gene to characterise microbial, algal and microfaunal communities in ponds. These are the first studies, to my knowledge, to use eDNA metabarcoding of the ITS2 gene region to sequence communities of green plants, green algae, and fungi in ponds.

The difference in spatial scale and grain - from high resolution sampling in a pond of 0.5ha, to sampling multiple ponds across a 3,600 km^2 landscape - the direct comparison of water and sediment samples, and year-round sampling is unique, and has elucidated various enlightening findings. These not only relate to the methodology of applying eDNA metabarcoding to ponds, but also the taxonomic composition and the ecological drivers of these communities. Furthermore, the comparison of eDNA metabarcoding with traditional pond monitoring techniques suggests new and exciting approaches for pond biomonitoring in the molecular age, as well as potential drawbacks.

This body of work is relevant and timely. Freshwater ponds are increasingly recognised as critical reservoirs of biodiversity and important for ecosystem functioning nationwide, as evidenced by the inclusion of freshwater ponds in the National Capital and Ecosystem Assessment programme (NCEA, Natural England 2022). However rapid, scalable molecular tools, and tools which assess health of all components of the pond ecosystem are currently absent.

Limitations

There were limitations to datasets generated in these studies. Time between sampling events was irregular for the seasonal study (Chapter 3) due to Government restrictions imposed by the Covid-19 pandemic, resulting in no peak spring sampling event (May) as planned. The Covid-19 pandemic also delayed progress with laboratory analysis, which had a knock-on effect on the timescale for the landscape study (Chapter 4). This landscape study was also limited by a relatively small sample size and number of ponds of each HSI category, due to constrictions on access, number of samplers (two) and time constraints.

Not all relevant environmental variables could be measured for all ponds due to funding constraints, for example, sediment physio-chemical properties. This would have been useful for a variety of reasons: to more clearly link taxa found in sediments to functions (e.g. methanotrophic bacteria and carbon content of sediments) and to discover other factors structuring communities in sediment. Additionally, I was unable to look in more detail at the prokaryotic functional assignment and composition of the study ponds due to time and cost constraints.

Longer reads of the ITS2 gene region would have aided in species-level identification of these sequences, but was not possible as the sequencing flow cells used (Illumina MiSeq v3) had an upper limit of 550bp, chosen to accommodate multiple barcodes of varying length (see Chapter 2, Table 1; Chapter 3, Table 1 and Chapter 4, Table 2).

Nevertheless, the large volume of eDNA and environmental data generated from this study, conducted over a relatively short timeframe, still permits impactful conclusions to be drawn.

Methodological findings

These studies have emphasised promising new directions for eDNA metabarcoding methods in ponds, as well as underlining complexities and drawbacks. My work has shown that applying eDNA metabarcoding in these environments is not always simple and there are a variety of technical challenges associated with sampling, molecular analysis and *in silico* analysis that need to be addressed to allow this method to be used reliably for biomonitoring.

My studies identified that the following issues are important when considering the methodology of pond eDNA metabarcoding studies:

1. Sample medium (sediment or water)
2. Season of sampling
3. Number of samples / total volume sampled per waterbody
4. Temperature and light levels and interaction with eDNA persistence
5. Marker gene
6. Taxonomic vs taxonomy-free approaches

Firstly, these studies have highlighted the differences in taxonomic composition, diversity and turnover between biological communities sequenced from pond sediments and from

pond water. No matter which taxa were studied, communities sequenced by eDNA metabarcoding were significantly different in water and sediment samples (Chapter 2: Figure 2.3 and Table 2.2; Chapter 3: Figure 3.5 and Figure 3.6; Chapter 4: Figure 4.5), although the differences were of greater magnitude in microbial communities (both prokaryotic and eukaryotic) than in multicellular organisms (Chapter 3: Tables 3.2 to 3.5).

Entire phyla, such as Desulfobacterota, Acidobacterota, Annelida and Nematoda were almost only found in sediment samples, whereas Actinobacteria, Cryptophyta and Rotifera were several times more abundant in water samples. This suggests that eDNA metabarcoding studies in ponds to date, which have almost exclusively sampled pond water, are failing to detect entire taxonomic and functional components of pond communities.

Secondly, the annual time-series study (Chapter 3) showcased that eDNA metabarcoding is sensitive to seasonal turnover in biological community composition in ponds. However, this fluctuation is much more pronounced in water samples than in sediment samples (Chapter 3: Figure 3.7). No matter which group was sampled, be it microbes, or multicellular animals, plants and fungi, the season of sampling was the strongest driver of changes in community composition in water samples, overriding the effects of abiotic pond conditions and explaining between 10 and 20 % of the variation in communities (Chapter 3: Tables 3.6-3.9). However, in sediment samples, the structuring effects of sample season and abiotic pond conditions were roughly equal, between 3 and 7 % (Chapter 3: Tables 3.6-3.9). No one season emerged as the optimal time to sample, as alpha diversity of different taxonomic groups peaked in different months, although cooler months tended to have higher alpha diversity (see below).

Furthermore, there was evidence that communities in sediment samples are more responsive to the effects of the surrounding landscape than in water samples: variation partitioning and dbRDAs indicated landscape variables were responsible for 10-20 % of variation in eukaryotic communities in sediment, with significant effects of arable and urban land cover and a greater number of ponds in the surrounding landscape (Chapter 4: Figure 4.7 and Figure 4.8). For prokaryotic communities in sediments, landscape variables were responsible for 17-30 % of variation, with significant effects of improved grassland, arable, and suburban land cover.

These three discoveries imply that it would be more advisable to sample pond sediments than pond water when conducting studies aimed at elucidating the enduring impacts of environmental factors on pond ecosystems, such as eutrophication or surrounding land use change. Sediment samples should be prioritised particularly if sampling is spread over a long

period (2 months and upwards). Conversely, for research investigating immediate responses and fluctuations, collecting pond water samples is a more suitable choice. My research shows that standard collection, eDNA extraction, amplification and sequencing protocols work well for pond sediments, and could be scaled to larger study areas. However, one drawback of my sediment sampling method (sample scoop) is the sample is less readily standardised. Sediment traps could be used to ensure only recent sediment is sampled, although this would entail two visits.

As predicted by Harper et al (2019), and as found in previous studies (Eichmiller et al 2014), communities revealed by eDNA metabarcoding water were highly heterogeneous within a single pond, with differences in composition at the class and order level between samples collected <20 m apart. Multiple and higher volume water samples are recommended for larger ponds, although this must be balanced with cost and practicality. More efficient methods of sampling water for eDNA metabarcoding, such as in-field filtration and filter preservation, are emerging constantly, which will aid this process (Spens et al 2016). Rarefaction curves could be used to determine the minimum number of samples required per pond. However, Chapter 4 demonstrated similarities in communities between ponds from different geographic areas, but similar conditions and quality (Chapter 4, Table 4.3, Figure 4.8). This suggests that eDNA metabarcoding a small volume of pond water or a few grams of sediment can be generalised to a whole pond, rather than being only a reflection of hyperlocal conditions.

Several of my findings suggest that variation in eDNA persistence rates, caused by differences in temperature and UV light levels between ponds or times of year may be obscuring ecological patterns. For instance, my findings of higher alpha diversity of microbes in the winter months (Chapter 3, Figures S3.6 and S3.7) and higher green plant and algae richness in shaded ponds (Chapter 4, Figure 4.6) are contrary to noted ecological patterns, but make sense if low light levels and temperatures were enabling greater eDNA persistence, which was in turn reflected in higher alpha diversity. It is well known that UV and temperature levels influence eDNA degradation rates (Barnes & Turner 2016), yet I have not seen this impact on eDNA persistence and diversity metrics acknowledged in other seasonal studies (e.g. Bista et al. 2017, Mikhailov et al. 2022).

This is a significant factor to consider when using eDNA metabarcoding for testing ecological hypotheses in ponds, and other freshwater environments as well. Previous eDNA studies should be re-examined to determine whether differences in abundance or diversity are due to an ecological reality or could be explained by eDNA preservation artefacts. In current and

future studies, I recommend that water temperature and shading are always measured and taken into consideration, particularly when comparing alpha diversities or studying phototrophic organisms.

These studies have shown that the ITS2 gene region is an effective marker for eDNA metabarcoding of green plant communities. Over 85 higher plant (Embryophyta*)* and over 280 green algal (Chlorophyta*)* taxa were identified to at least genus level, with a mean of 24.3 taxa identified per sample. However, the overlap with traditional in-pond macrophyte surveys was low, approximately 10 % at species level (Chapter 2 and Chapter 3). It appeared that if there were multiple species of the same genus present, eDNA metabarcoding struggled to resolve these, and many species, both common and rare, were not detected. Future efforts should focus on expanding reference databases for macrophytes, and making them more region-specific, as has recently been attempted in Quaresma et al. (2024), although coverage at species level is still short of 100 %. Longer length DNA barcodes could be used to aid in the resolution of identification (Fahner et al. 2016).

A consistent pattern emerged from all three studies: eDNA from terrestrial plants made up 30 % of total ITS2 reads in sediment samples, and 70 % of total reads in water samples. This highlights the strong terrestrial influence on small freshwater environments such as ponds, and the ability for eDNA to travel over long distances and between different environments. Future metabarcoding studies, whether using the ITS2 marker or others, should carefully consider the influence of allochthonous eDNA on their conclusions. Future research could investigate using eDNA metabarcoding of pond water to passively sample plants in the surrounding landscape, or further investigate the use of Chlorophyta taxa as an indicators of pond quality. These findings further emphasise the benefits of sampling sediments alongside, or instead of, water for eDNA surveys of pond environments.

Finally, my results bolster confidence in current standard bioinformatic processing procedures. Contrary to previous findings (Ionescu et al 2022, Machler et al 2020), using eDNA read abundances as a proxy for species abundance did not greatly alter conclusions drawn about community composition compared to only using presence-absence data (Chapter 2, Table 2.2). Similarly, filtering by confidence of taxonomic assignment or using an unfiltered dataset had little effect on overall patterns of community structure (Chapter 2, Table 2.2). This endorses the use of taxonomy-free approaches to using eDNA metabarcoding data for biological monitoring in ponds (Cordier et al. 2020, Pawlowski et al. 2018). Future efforts could compare taxonomically assigned and taxonomy-free datasets

across a metric of known environmental gradients (e.g. Great Crested Newt HSI levels) for a larger dataset.

Taxonomic and bioindicator findings

Sequencing the prokaryotic communities across the 50+ ponds in these combined studies has revealed a similar pattern of taxonomic composition at the phylum level, and much in common with other freshwater environments, with the most abundant phyla belonging to Proteobacteria and Bacertoidota and, in water samples, Actinobacteria. There was a consistently low abundance (< 5 % of all reads) of Cyanobacteria.

eDNA metabarcoding across the three studies found taxa involved in nutrient cycling and biotic interactions, from Desulfobacterota (sulphate-reducing bacteria phylum, also commonly present in wetland soils), to methanogenic and methane oxidising genera, to polyphosphate accumulating organisms and potentially pathogenic organisms (Chapters 2, 3 and 4). This highlights the crucial importance of bacteria and archaea to the ecosystem functioning of ponds, such as nutrient retention, and generation or mitigation of carbon emissions.

A compelling future avenue of research would be to use metagenomics or metatranscriptomics to describe the functional profiles of pond environments (Cordier et al 2020). This could help elucidate current, societally relevant research questions such as whether, and under what conditions, ponds are methane sources or sinks (Rosentrater et al 2021), their role in nitrogen cycling (Cai et al. 2022), pollutant remediation (Tai et al. 2020) and the development of antimicrobial resistance (Nnadozie & Odume 2019).

The landscape-scale study draws attention to the immense potential for bacteria and archaea to be bioindicator organisms in pond environments. Whilst consistently similar across ponds at the phylum level, at the genus level, bacteria community composition altered according to water physio-chemistry and surrounding land use (constrained ordinations and variation partitioning, Chapter 4, Figures 4.7 and 4.8) and pond overall quality as calculated by the Habitat Suitability Index for Great Crested Newts (unconstrained ordination, Chapter 4, Table 4.3). Whether alpha diversity (Chapter 4, Figure 4.6), indicator taxa or beta diversity was used as a metric, prokaryotes showed evidence of anthropogenic impact on ponds.

This reinforces previous findings of prokaryotes indicating human-induced pressures in other environments, and provides support for calls to shift the research focus from macro-organisms to micro-organisms as bioindicators (Sagova-Mareckova et al 2021). That these patterns emerged from a relatively low total sample size (due to the constraints mentioned above) provides strong justification for future efforts to develop a prokaryotic pond health indicator at a wider scale (e.g., nationwide).

By contrast, the sequencing of eukaryotic communities using the 18S rRNA marker gene revealed a lack of consistency in taxonomic composition, even at the phylum level, between different ponds. Similar phyla were sequenced compared to other studies using the 18S rRNA marker gene in lakes, such as Ochrophyta, Ciliopora, Cercozoa, Arthropoda, Nematoda and Annelida. However, Chlorophyta reads appeared to be more abundant in my research than in eDNA metabarcoding studies of lake communities, suggesting that ponds, due to their shallow depth and high light penetration, are important habitats for this taxon. Cryptophyta was also present in water samples and absent from sediment samples from the same location, similar to prior research in ponds conducted by Ionescu et al (2022).

Community composition and diversity detected by 18S rRNA metabarcoding did not reflect overall pond ecological quality, save for communities in pond sediments being structured by surrounding land use, and the presence of some phototrophic indicator taxa of “Excellent” HSI ponds *(*Sphaeropleales and Peridiniales*).* Unlike the prokaryote community, using the whole eukaryote community as a bioindicator of anthropogenic impact on ponds is not indicated in my data. It may be that biotic interactions such as competition and predation are more important in structuring eukaryote communities. Alternatively, it may be that responses are evident in lower taxonomic levels (e.g. Bacillariophyta, Cryptophyta, Rotifera*)* or different functional groups (e.g. phytoplankton, zooplankton, meiobenthos) which are obscured in combination with other groups. Subsequent research initiatives could focus on responses to gradients of disturbance of some of these taxa, although prior research has not found strong responses in these groups (e.g. Ionescu et al 2022). There is also a potential role for different statistical methods to identify indicator taxa at the species level, such as machine learning.

Ecological findings

My research has demonstrated the success of using eDNA metabarcoding to investigate ecological hypotheses for pond ecosystems at the landscape level. Prokaryote communities

in the water and sediments of ponds were structured by pond water physio-chemistry (particularly total conductivity and total phosphorus), and prokaryote communities in sediment additionally responded to the surrounding land use, particularly agricultural (arable and improved grassland) cover, and suburban land use. Prokaryote communities showed no evidence of spatial structuring at this scale, suggesting an absence of dispersal limitation.

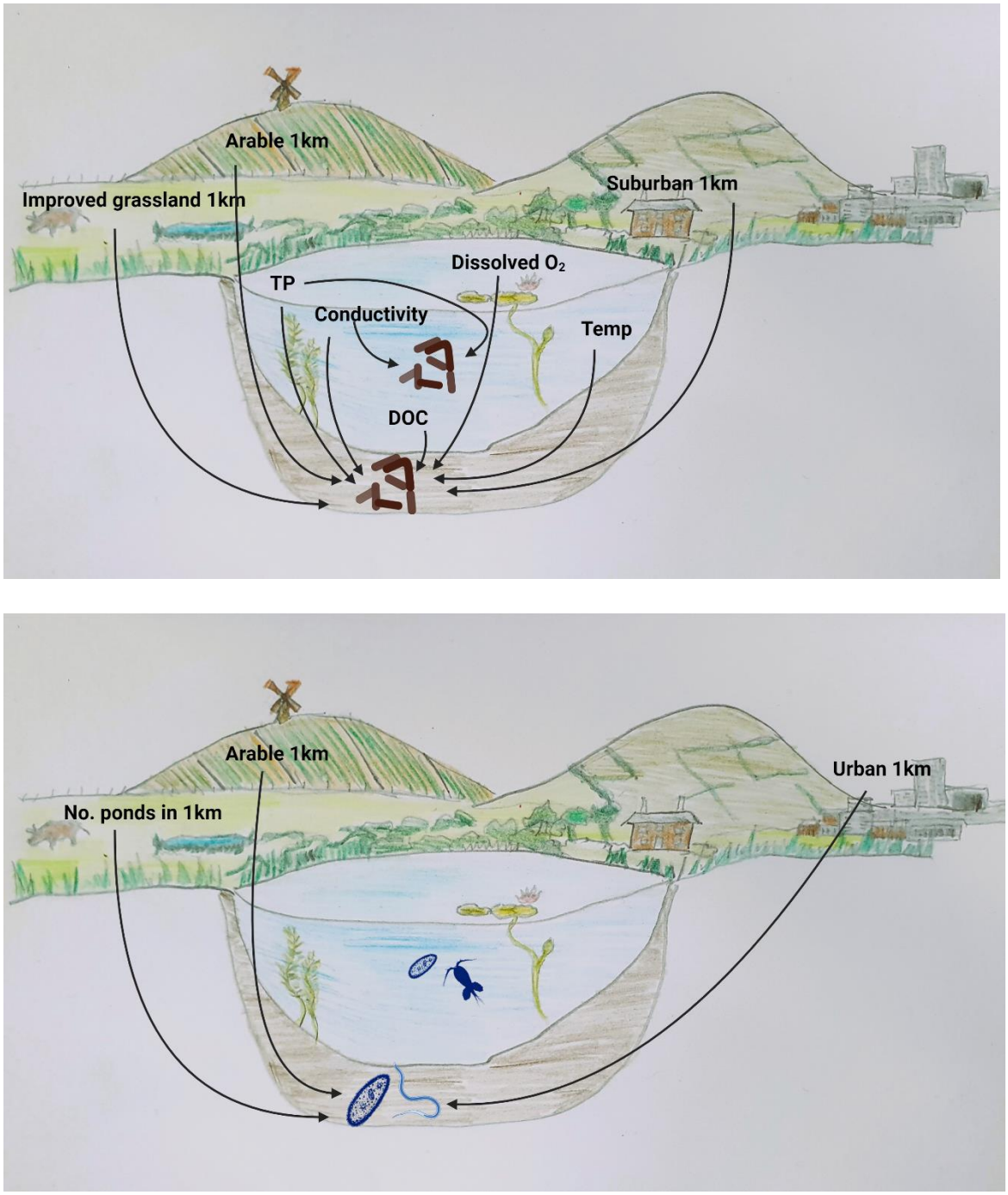


Figure 5.1: Schematic diagrams to demonstrate the structuring influence of environmental factors on (top) prokaryotic communities in pond water and sediments and (bottom)

eukaryotic communities in pond sediments (n=31). Arrows indicate that the factor was found to have a significant structuring effect in a constrained ordination (dbRDA). Green plant and algae communities did not display environmental structuring but were spatially structured. For full results, see Chapter 4. "TP" = total phosphorus (µg/l), "DOC" = dissolved organic carbon (mg/l), "Temp" = Temperature (°C).

This study found evidence of eukaryote communities in pond sediments being environmentally filtered according to the land use in the 1 km surrounding the pond, particularly the presence of arable and urban land cover. There was some evidence that freshwater connectivity also influences these communities, as the number of ponds within a 1 km radius was significantly correlated with gradients of community dissimilarity. There was no evidence of spatial structuring or dispersal limitation. In eukaryote communities in pond water, there was no evidence of either environmental filtering or dispersal limitation.

Finally, green plants and algal communities were not structured according to environmental variables, either at the individual pond (chemical, physical) or landscape level. However, these communities were spatially structured, suggesting that dispersal limitation is impacting both macrophytes and green algal distribution in these ponds.

Future research efforts could apply similar techniques across different landscape types (e.g., upland versus lowland, semi-natural vs highly altered), greater spatial extents and with a greater number of waterbodies.

Conclusions

In summary, the collection of studies presented here represents a significant advancement in the application of environmental DNA (eDNA) metabarcoding techniques to the study of freshwater ponds. These studies have illuminated various hitherto unseen aspects of pond ecosystems, shedding light on both methodological considerations and ecological insights.

Methodologically, the research has revealed the complexities of applying eDNA metabarcoding in pond environments. It underscores the importance of considering the choice of sample medium (water, sediment, or both), and the influence of seasonal fluctuations on community composition. The findings also suggest that the influence of local abiotic conditions on eDNA persistence, such as temperature and shading, should always be considered when applying eDNA metabarcoding to ponds.

The studies have demonstrated the potential of the ITS2 gene region as an effective marker for characterizing green plant communities in ponds, though challenges remain in resolving certain taxonomic distinctions. Moreover, they emphasize the strong influence of the surrounding terrestrial landscape on the detection of plant communities in small waterbodies.

Taxonomically and ecologically, these studies have uncovered valuable insights into the composition and dynamics of prokaryotic and eukaryotic communities in ponds. Prokaryotes have emerged as potential bioindicator organisms sensitive to environmental changes, indicating the anthropogenic impacts on pond health. In contrast, eukaryote communities appear to be less clearly influenced by the pond and surrounding environment, making them less suitable as bioindicators of anthropogenic effects.

My research demonstrates the influence of environmental factors, water physio-chemistry, and land use on prokaryote and eukaryote communities, along with the impact of spatial structuring and dispersal limitation on green plant and algal communities. Together, these findings emphasise that conservation of these unique and ecologically important habitats requires a variety of approaches: overcoming dispersal limitation by creation of new ponds and human mediated colonisation of macrophyte and algae species, landscape-scale measures such as protecting large areas of natural or semi-natural habitat, and measures to improve water quality of individual ponds, such as water-friendly farming practises.

In sum, this body of work not only advances our understanding of pond ecosystems but also highlights the challenges and potential of eDNA metabarcoding in small freshwater environments. It serves as a foundation for future research, monitoring and conservation efforts aimed at safeguarding these beautiful and precious ecosystems.

References

Adams, C. I. M., Knapp, M., Gemmell, N. J., Jeunen, G.-J., Bunce, M., Lamare, M. D., & Taylor, H. R. (2019). Beyond Biodiversity: Can Environmental DNA (eDNA) Cut It as a Population Genetics Tool? *Genes*, *10*(3), 192. https://doi.org/10.3390/genes10030192

Adams, D. G. (2002). Cyanobacteria in Symbiosis with Hornworts and Liverworts. *Cyanobacteria in Symbiosis*, 117–135. https://doi.org/10.1007/0-306-48005-0_7

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., Berney, C., Brown, M. W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., … Zhang, Q. (2019). Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *Journal of Eukaryotic Microbiology*, *66*(1), 4–119. https://doi.org/10.1111/jeu.12691

Adyasari, D., Hassenrück, C., Montiel, D., & Dimova, N. (2020). Microbial community composition across a coastal hydrological system affected by submarine groundwater discharge (SGD). *PLOS ONE*, *15*(6), e0235235. https://doi.org/10.1371/journal.pone.0235235

Ahmed, W., Payyappat, S., Cassidy, M., & Besley, C. (2019). Enhanced insights from human and animal host-associated molecular marker genes in a freshwater lake receiving wet weather overflows. *Scientific Reports 2019 9:1*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-48682-4

Akasaka, M., & Takamura, N. (2012). Hydrologic connection between ponds positively affects macrophyte α and γ diversity but negatively affects β diversity. *Ecology*, *93*(5), 967–973. https://doi.org/10.1890/11-0879.1

Alahuhta, J., Johnson, L. B., Olker, J., & Heino, J. (2014). Species sorting determines variation in the community composition of common and rare macrophytes at various spatial extents. *Ecological Complexity*, *20*, 61–68. https://doi.org/10.1016/j.ecocom.2014.08.003

Alsos, I. G., Lammers, Y., Yoccoz, N. G., Jørgensen, T., Sjögren, P., Gielly, L., & Edwards, M. E. (2018). Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLoS ONE*, *13*(4), e0195403. https://doi.org/10.1371/journal.pone.0195403

An, J., Liu, C., Wang, Q., Yao, M., Rui, J., Zhang, S., & Li, X. (2019a). Soil bacterial community structure in Chinese wetlands. *Geoderma*, *337*, 290–299. https://doi.org/10.1016/j.geoderma.2018.09.035

An, J., Liu, C., Wang, Q., Yao, M., Rui, J., Zhang, S., & Li, X. (2019b). Soil bacterial community structure in Chinese wetlands. *Geoderma*, *337*, 290–299. https://doi.org/10.1016/j.geoderma.2018.09.035

Andres, K. J., Sethi, S. A., Lodge, D. M., & Andrés, J. (2021). Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. *Molecular Ecology*, *30*(3), 685–697. https://doi.org/10.1111/mec.15765

Angélibert, S., Marty, P., Céréghino, R., & Giani, N. (2004). Seasonal variations in the physical and chemical characteristics of ponds: Implications for biodiversity conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *14*(5), 439–456. https://doi.org/10.1002/aqc.616

Anglès d'Auriac, M. B., Strand, D. A., Mjelde, M., Demars, B. O. L., & Thaulow, J. (2019). Detection of an invasive aquatic plant in natural water bodies using environmental DNA. *PLOS ONE*, *14*(7), e0219700. https://doi.org/10.1371/journal.pone.0219700

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*(6), 1231–1242. https://doi.org/10.1111/1755-0998.12668

ARG. (2010). *Amphibian and Reptile Groups of the United Kingdom ARG UK Advice Note 5 Great Crested Newt Habitat Suitability Index*. www.arguk.org

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. In *Molecular Ecology* (Vol. 21, Issue 8, pp. 2039–2044). https://doi.org/10.1111/j.1365-294X.2012.05519.x

Banchi, E., Ametrano, C. G., Stanković, D., Verardo, P., Moretti, O., Gabrielli, F., Lazzarin, S., Borney, M. F., Tassan, F., Tretiach, M., Pallavicini, A., & Muggia, L. (2018). DNA metabarcoding uncovers fungal diversity of mixed airborne samples in Italy. *PLOS ONE*, *13*(3), e0194489. https://doi.org/10.1371/journal.pone.0194489

Banerji, A., Bagley, M., Elk, M., Pilgrim, E., Martinson, J., & Santo Domingo, J. (2018a). Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. *Hydrobiologia*, *818*(1), 71–86. https://doi.org/10.1007/s10750-018-3593-0

Banerji, A., Bagley, M., Elk, M., Pilgrim, E., Martinson, J., & Santo Domingo, J. (2018b). Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. *Hydrobiologia*, *818*(1), 71–86. https://doi.org/10.1007/s10750-018-3593-0

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. In *Conservation Genetics* (Vol. 17, Issue 1, pp. 1–17). Springer Netherlands. https://doi.org/10.1007/s10592-015-0775-4

Baudelet, P. H., Ricochon, G., Linder, M., & Muniglia, L. (2017). A new insight into cell walls of Chlorophyta. *Algal Research*, *25*, 333–371. https://doi.org/10.1016/J.ALGAL.2017.04.008

Beattie, R. E., Bandla, A., Swarup, S., & Hristova, K. R. (2020). Freshwater Sediment Microbial Communities Are Not Resilient to Disturbance From Agricultural Land Runoff. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.539921

Beck, M. (2022). *ggord: Ordination Plots with ggplot2. R package 1.1.7*.

Belle, C. C., Stoeckle, B. C., & Geist, J. (2019). Taxonomic and geographical representation of freshwater environmental DNA research in aquatic conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*. https://doi.org/10.1002/aqc.3208

Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodiversity and Conservation*, *29*(7), 2089–2121. https://doi.org/10.1007/s10531-020-01980-0

Bessey, C., Neil Jarman, S., Simpson, T., Miller, H., Stewart, T., Kenneth Keesing, J., & Berry, O. (2021). Passive eDNA collection enhances aquatic biodiversity analysis. *Communications Biology*, *4*(1), 236. https://doi.org/10.1038/s42003-021-01760-8

Bichsel, D., De Marco, P., Bispo, A. Â., Ilg, C., Dias-Silva, K., Vieira, T. B., Correa, C. C., & Oertli, B. (2016). Water quality of rural ponds in the extensive agricultural landscape of the Cerrado (Brazil). *Limnology*, *17*(3), 239–246. https://doi.org/10.1007/s10201-016-0478-7

Biggs, J., Corfield, A., & Walker, D. (1994). New Approaches to the Management of Ponds. *British Wildlife*, *5*(5). https://www.researchgate.net/publication/237468011

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

Biggs J, Ewald N, Valentini A, Gaboriaud C, Griffiths RA, Foster J, Wilkinson J, Arnett A, W. P. and D. F., Biggs J, Ewald N, Valentini A, Gaboriaud C, Griffiths RA, Foster J, Wilkinson J, A. A., & F, W. P. and D. (2014). *Analytical and methodological development for improved surveillance of the Great Crested Newt. Appendix 5. Technical advice note for field and laboratory sampling of great crested newt (Triturus cristatus) environmental DNA*.

Biggs, J., Fox, G., Nicolet, P., Walker, D., Whitfield, M., & Williams, P. (1998). *A guide to the methods of the National Pond Survey*.

Biggs, J., von Fumetti, S., & Kelly-Quinn, M. (2017). The importance of small waterbodies for biodiversity and ecosystem services: implications for policy makers. In *Hydrobiologia* (Vol. 793, Issue 1, pp. 3–39). Springer International Publishing. https://doi.org/10.1007/s10750-016-3007-0

Biggs, J., & Williams, P. (2023a). *Ponds, Pools and Puddles*. Harper Collins New Naturalist.

Biggs, J., & Williams, P. (2023b). *Ponds, Pools and Puddles: Vol. New Naturalist 148*. HarperCollins UK.

Biggs, J., Williams, P., Whitfield, M., Nicolet, P., Brown, C., Hollis, J., Arnold, D., & Pepper, T. (2007). The freshwater biota of British agricultural landscapes and their sensitivity to pesticides. *Agriculture, Ecosystems and Environment*, *122*(2), 137–148. https://doi.org/10.1016/j.agee.2006.11.013

Biggs, J., Williams, P., Whitfield, M., Nicolet, P., & Weatherby, A. (2005). 15 years of pond assessment in Britain: results and lessons learned from the work of Pond Conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*. https://doi.org/10.1002/aqc.745

Bilton, D. T., Freeland, J. R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. In *Annual Review of Ecology and Systematics* (Vol. 32, pp. 159–181). https://doi.org/10.1146/annurev.ecolsys.32.081501.114016

Binh, C. T. T., Tong, T., Gaillard, J.-F., Gray, K. A., & Kelly, J. J. (2014). Acute Effects of TiO2 Nanomaterials on the Viability and Taxonomic Composition of Aquatic Bacterial Communities Assessed via High-Throughput Screening and Next Generation Sequencing. *PLoS ONE*, *9*(8), e106280. https://doi.org/10.1371/journal.pone.0106280

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Christmas, M., & Creer, S. (2017a). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, *8*(1), 14087. https://doi.org/10.1038/ncomms14087

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

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Christmas, M., & Creer, S. (2017c). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, *8*(1), 14087. https://doi.org/10.1038/ncomms14087

Boessenkool, S., Mcglynn, G., Epp, L. S., Taylor, D., Pimentel, M., Gizaw, A., Nemomissa, S., Brochmann, C., & Popp, M. (2014). Use of ancient sedimentary DNA as a novel conservation tool for high-altitude tropical biodiversity. *Conservation Biology*, *28*(2), 446–455. https://doi.org/10.1111/cobi.12195

Bolgovics, Á., B-Béres, V., Várbíró, G., Krasznai-K, Á., Ács, É., Kiss, K. T., Borics, G., Krasznai-K, E. Á., Ács, É., Kiss, K. T., & Borics, G. (2019). Groups of small lakes maintain larger microalgal diversity than large ones. *Science of the Total Environment*, *678*, 162–172. https://doi.org/10.1016/j.scitotenv.2019.04.309

Bolpagni, R., Laini, A., Buldrini, F., Ziccardi, G., Soana, E., Pezzi, G., Chiarucci, A., Lipreti, E., Armiraglio, S., & Nascimbene, J. (2020). Habitat morphology and connectivity better predict hydrophyte and wetland plant richness than land-use intensity in overexploited watersheds: evidence from the Po plain (northern Italy). *Landscape Ecology*, *35*(8), 1827–1839. https://doi.org/10.1007/s10980-020-01060-2

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bongaarts, J. (2019). IPBES, 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform

on Biodiversity and Ecosystem Services. *Population and Development Review*, *45*(3), 680–681. https://doi.org/10.1111/padr.12283

Boothby, J. (1997). Pond conservation: Towards a delineation of pondscape. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *7*(2), 127–132. https://doi.org/10.1002/(SICI)1099-0755(199706)7:2<127::AID-AQC224>3.0.CO;2-6

Bowes, M. J., Armstrong, L. K., Harman, S. A., Wickham, H. D., Nicholls, D. J. E., Scarlett, P. M., Roberts, C., Jarvie, H. P., Old, G. H., Gozzard, E., Bachiller-Jareno, N., & Read, D. S. (2018). Weekly water quality monitoring data for the River Thames (UK) and its major tributaries (2009–2013): the Thames Initiative research platform. *Earth System Science Data*, *10*(3), 1637–1653. https://doi.org/10.5194/essd-10-1637-2018

Bradley, I. M., Pinto, A. J., & Guest, J. S. (2016). Design and evaluation of illumina MiSeq-compatible, 18S rRNA gene-specific primers for improved characterization of mixed phototrophic communities. *Applied and Environmental Microbiology*, *82*(19), 5878–5891. https://doi.org/10.1128/AEM.01630-16

Brantschen, J., Blackman, R. C., Walser, J. C., & Altermatt, F. (2021). Environmental dna gives comparable results to morphology-based indices of macroinvertebrates in a large-scale ecological assessment. *PLoS ONE*, *16*(9 September). https://doi.org/10.1371/JOURNAL.PONE.0257510

Bräuer, S. L., Cadillo-Quiroz, H., Yashiro, E., Yavitt, J. B., & Zinder, S. H. (2006). Isolation of a novel acidiphilic methanogen from an acidic peat bog. *Nature*, *442*(7099), 192–194. https://doi.org/10.1038/nature04810

Brown, L. E., Khamis, K., Wilkes, M., Blaen, P., Brittain, J. E., Carrivick, J. L., Fell, S., Friberg, N., Füreder, L., Gislason, G. M., Hainie, S., Hannah, D. M., James, W. H. M., Lencioni, V., Olafsson, J. S., Robinson, C. T., Saltveit, S. J., Thompson, C., & Milner, A. M. (2017). Functional diversity and community assembly of river invertebrates show globally consistent responses to decreasing glacier cover. *Nature Ecology & Evolution*, *2*(2), 325–333. https://doi.org/10.1038/s41559-017-0426-x

Buchner, D., & Leese, F. (2020). BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, *4*. https://doi.org/10.3897/mbmg.4.53535

Bukin, Y. S., Galachyants, Y. P., Morozov, I. V., Bukin, S. V., Zakharenko, A. S., & Zemskaya, T. I. (2019). The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific Data 2019 6:1*, *6*(1), 1–14. https://doi.org/10.1038/sdata.2019.7

Burns, L., Mordue, S., & al Fulaij, N. (2023). *State of Nature 2023*.

Burton, R. (1977). *Ponds*. David & Charles.

Bush, A., Compson, Z. G., Monk, W. A., Porter, T. M., Steeves, R., Emilson, E., Gagne, N., Hajibabaei, M., Roy, M., & Baird, D. J. (2019). Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates. In *Frontiers in Ecology and Evolution* (Vol. 7, p. 434). Frontiers Media S.A. https://doi.org/10.3389/fevo.2019.00434

Buxton, A. S., Tracey, H., & Downs, N. C. (2021). How reliable is the habitat suitability index as a predictor of great crested newt presence or absence? *Herpetological Journal*, *31*(2), 111–117. https://doi.org/10.33256/31.2.111117

Cai, M., Li, S., Ye, F., Hong, Y., Lü, M., Op den Camp, H. J. M., & Wang, Y. (2022). Artificial ponds as hotspots of nitrogen removal in agricultural watershed. *Biogeochemistry*, *159*(3), 283–301. https://doi.org/10.1007/s10533-022-00928-6

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

Capo, E., Debroas, D., Arnaud, F., Guillemot, T., Bichet, V., Millet, L., Gauthier, E., Massa, C., Develle, A.-L., Pignol, C., Lejzerowicz, F., & Domaizon, I. (2016). Long-term dynamics in microbial eukaryotes communities: a palaeolimnological view based on sedimentary DNA. *Molecular Ecology*, *25*(23), 5925–5943. https://doi.org/10.1111/mec.13893

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal 2012 6:8*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

Carew, M. E., Kellar, C. R., Pettigrove, V. J., & Hoffmann, A. A. (2018). Can high-throughput sequencing detect macroinvertebrate diversity for routine monitoring of an urban river? *Ecological Indicators*, *85*, 440–450. https://doi.org/10.1016/j.ecolind.2017.11.002

Carey, P. D. (2008). *Countryside Survey: UK Headline Messages from 2007*. www.countrysidesurvey.org.uk

Catchment Based Approach (CaBA). (2018). *Biodiversity Pack Habitat Guide: Ponds*. https://catchmentbasedapproach.org/wp-content/uploads/2018/09/CaBA-Biodiversity-Pack-Ponds.pdf

Celewicz, S., Kozak, A., & Kuczyńska-Kippen, N. (2022). Chlorophytes response to habitat complexity and human disturbance in the catchment of small and shallow aquatic systems. *Scientific Reports*, *12*(1), 13050. https://doi.org/10.1038/s41598-022-17093-3

Céréghino, R., Biggs, J., Oertli, B., & Declerck, S. (2007). The ecology of European ponds: defining the characteristics of a neglected freshwater habitat. In *Pond Conservation in Europe* (pp. 1–6). Springer Netherlands. https://doi.org/10.1007/978-90-481-9088-1_1

Céréghino, R., Boix, D., Cauchie, H. M., Martens, K., & Oertli, B. (2014). The ecological role of ponds in a changing world. In *Hydrobiologia* (Vol. 723, Issue 1, pp. 1–6). https://doi.org/10.1007/s10750-013-1719-y

Chalde, T., Nardi, C. F., & Fernández, D. A. (2019). Early warning: detection of exotic coho salmon (Oncorhynchus kisutch) by environmental DNA and evidence of establishment at the extreme south of Patagonia. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(12), 2343–2349. https://doi.org/10.1139/cjfas-2018-0271

Chang, C., Gao, L., Wei, J., Ma, N., He, Q., Pan, B., & Li, M. (2021). Spatial and environmental factors contributing to phytoplankton biogeography and biodiversity in mountain ponds across a large geographic area. *Aquatic Ecology*, *55*(2), 721–735. https://doi.org/10.1007/s10452-021-09857-2

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C., & Díaz, S. (2000). Consequences of changing biodiversity. In *Nature* (Vol. 405, Issue 6783, pp. 234–242). Nature Publishing Group. https://doi.org/10.1038/35012241

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., & Leon, C. (2010). Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. *PLOS ONE*, *5*(1), e8613. https://doi.org/10.1371/JOURNAL.PONE.0008613

Chen, W., Wilkes, G., Khan, I. U. H., Pintar, K. D. M., Thomas, J. L., Lévesque, C. A., Chapados, J. T., Topp, E., & Lapen, D. R. (2018). Aquatic bacterial communities associated with land use and environmental factors in agricultural landscapes using a metabarcoding approach. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02301

Chopyk, J., Nasko, D. J., Allard, S., Bui, A., Pop, M., Mongodin, E. F., & Sapkota, A. R. (2020). Seasonal dynamics in taxonomy and function within bacterial and viral metagenomic assemblages recovered from a freshwater agricultural pond. *Environmental Microbiome*, *15*(1), 18. https://doi.org/10.1186/s40793-020-00365-8

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J.-C., Bonin, A., Taberlet, P., & Pont, D. (2016). Spatial Representativeness of Environmental DNA Metabarcoding Signal for Fish Biodiversity Assessment in a Natural Freshwater System. *PLOS ONE*, *11*(6), e0157366. https://doi.org/10.1371/journal.pone.0157366

Clarke, S. J. (2009). Adapting to climate change: implications for freshwater biodiversity and management in the UK. *Freshwater Reviews*, *2*, 51–64. https://doi.org/10.1608/FRJ-2.1.3

Coghlan, S. A., Shafer, A. B. A., & Freeland, J. R. (2020). Development of an environmental DNA metabarcoding assay for aquatic vascular plant communities. *Environmental DNA*, edn3.120. https://doi.org/10.1002/edn3.120

Community Ecology Package [R package vegan version 2.6-4]. (2022a). https://CRAN.R-project.org/package=vegan

Community Ecology Package [R package vegan version 2.6-4]. (2022b). https://CRAN.R-project.org/package=vegan

Comte, J., Monier, A., Crevecoeur, S., Lovejoy, C., & Vincent, W. F. (2016). Microbial biogeography of permafrost thaw ponds across the changing northern landscape. *Ecography*, *39*(7), 609–618. https://doi.org/10.1111/ecog.01667

Cordier, T., Sáez, L. A., Apotheloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., Chariton, A., Creer, S., Fruhe, L., Keck, F., Keeley, N., Laroche, O., Leese, F., Pochon, X., Stoeck, T., Pawlowski, J., & Lanzén, A. (2020). Ecosystems Monitoring Powered by

Environmental Genomics: A Review of Current Strategies with An Implementation Roadmap. *Preprints.Org*, *January*. https://doi.org/10.20944/PREPRINTS202001.0278.V1

Cosentino, B. J., Marsh, D. M., Jones, K. S., Apodaca, J. J., Bates, C., Beach, J., Beard, K. H., Becklin, K., Bell, J. M., Crockett, C., Fawson, G., Fjelsted, J., Forys, E. A., Genet, K. S., Grover, M., Holmes, J., Indeck, K., Karraker, N. E., Kilpatrick, E. S., … Willey, A. (2014). Citizen science reveals widespread negative effects of roads on amphibian distributions. *Biological Conservation*, *180*, 31–38. https://doi.org/10.1016/j.biocon.2014.09.027

Coste, M., Boutry, S., Tison-Rosebery, J., & Delmas, F. (2009). Improvements of the Biological Diatom Index (BDI): Description and efficiency of the new version (BDI-2006). *Ecological Indicators*, *9*(4), 621–650. https://doi.org/10.1016/j.ecolind.2008.06.003

Cottenie, K., Michels, E., Nuytten, N., & De Meester, L. (2003). ZOOPLANKTON METACOMMUNITY STRUCTURE: REGIONAL VS. LOCAL PROCESSES IN HIGHLY INTERCONNECTED PONDS. *Ecology*, *84*(4), 991–1000. https://doi.org/https://doi.org/10.1890/0012-9658(2003)084[0991:ZMSRVL]2.0.CO;2

Crevecoeur, S., Vincent, W. F., Comte, J., & Lovejoy, C. (2015). Bacterial community structure across environmental gradients in permafrost thaw ponds: methanotroph-rich ecosystems. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00192

Cruaud, P., Vigneron, A., Dorea, C. C., Rodriguez, M. J., & Charette, S. J. (2020). Rapid Changes in Microbial Community Structures along a Meandering River. *Microorganisms*, *8*(11), 1631. https://doi.org/10.3390/microorganisms8111631

Cruaud, P., Vigneron, A., Fradette, M., Dorea, C. C., Culley, A. I., Rodriguez, M. J., & Charette, S. J. (2020). Annual bacterial community cycle in a seasonally ice-covered river reflects environmental and climatic conditions. *Limnology and Oceanography*, *65*(S1). https://doi.org/10.1002/lno.11130

Cruaud, P., Vigneron, A., Fradette, M.-S., Dorea, C. C., Culley, A. I., Rodriguez, M. J., & Charette, S. J. (2019). Annual Protist Community Dynamics in a Freshwater Ecosystem Undergoing Contrasted Climatic Conditions: The Saint-Charles River (Canada). *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.02359

CSABAI, Z., BODA, P., BERNÁTH, B., KRISKA, G., & HORVÁTH, G. (2006). A 'polarisation sun-dial' dictates the optimal time of day for dispersal by flying aquatic insects. *Freshwater Biology*, *51*(7), 1341–1350. https://doi.org/10.1111/j.1365-2427.2006.01576.x

Cuenca-Cambronero, M., Blicharska, M., Perrin, J.-A., Davidson, T. A., Oertli, B., Lago, M., Beklioglu, M., Meerhoff, M., Arim, M., Teixeira, J., De Meester, L., Biggs, J., Robin, J., Martin, B., Greaves, H. M., Sayer, C. D., Lemmens, P., Boix, D., Mehner, T., … Brucet, S. (2023). Challenges and opportunities in the use of ponds and pondscapes as Nature-based Solutions. *Hydrobiologia*, *850*(15), 3257–3271. https://doi.org/10.1007/s10750-023-05149-y

Cui, M., Ma, A., Qi, H., Zhuang, X., & Zhuang, G. (2015). Anaerobic oxidation of methane: an "active" microbial process. *MicrobiologyOpen*, *4*(1), 1–11. https://doi.org/10.1002/mbo3.232

Curtis, A. N., Larson, E. R., & Davis, M. A. (2021). Field storage of water samples affects measured environmental DNA concentration and detection. *Limnology*, *22*(1), 1–4. https://doi.org/10.1007/s10201-020-00634-y

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life*. John Murray.

Darwin, C. R. (1882). On the dispersal of freshwater bivalves. *Nature. A Weekly Illustrated Journal of Science*, *25*(6 April), 529–530.

Davies, B., Biggs, J., Williams, P., Whitfield, M., Nicolet, P., Sear, D., Bray, S., & Maund, S. (2008). Comparative biodiversity of aquatic habitats in the European agricultural landscape. *Agriculture, Ecosystems and Environment*, *125*(1–4), 1–8. https://doi.org/10.1016/j.agee.2007.10.006

Davies, Z. G., Fuller, R. A., Loram, A., Irvine, K. N., Sims, V., & Gaston, K. J. (2009). A national scale inventory of resource provision for biodiversity within domestic gardens. *Biological Conservation*, *142*(4), 761–771. https://doi.org/10.1016/j.biocon.2008.12.016

Davy, C. M., Kidd, A. G., & Wilson, C. C. (2015). Development and Validation of Environmental DNA (eDNA) Markers for Detection of Freshwater Turtles. *PLOS ONE*, *10*(7), e0130965. https://doi.org/10.1371/journal.pone.0130965

Davy-Bowker, J., Clarke, R. T., Johnson, R. K., Kokes, J., Murphy, J. F., & Zahrádková, S. (2006). A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. *Hydrobiologia*, *566*(1), 91–105. https://doi.org/10.1007/s10750-006-0068-5

De Bie, T., Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Gucht, K., Wichelen, J., Vyverman, W., & Declerck, S. A. J. (2012). Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, *15*(7), 740–747. https://doi.org/10.1111/j.1461-0248.2012.01794.x

De Meester, L., Declerck, S., Stoks, R., Louette, G., Van De Meutter, F., De Bie, T., Michels, E., & Brendonck, L. (2005). Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *15*(6), 715–725. https://doi.org/10.1002/aqc.748

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome *c* oxidase subunit I marker: not a perfect match. *Biology Letters*, *10*(9), 20140562. https://doi.org/10.1098/rsbl.2014.0562

Debroas, D., Domaizon, I., Humbert, J.-F., Jardillier, L., Lepère, C., Oudart, A., & Taïb, N. (2017). Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbiology Ecology*, *93*(4). https://doi.org/10.1093/femsec/fix023

Declerck, S., De Bie, T., Ercken, D., Hampel, H., Schrijvers, S., Van Wichelen, J., Gillard, V., Mandiki, R., Losson, B., Bauwens, D., Keijers, S., Vyverman, W., Goddeeris, B., De meester,

L., Brendonck, L., & Martens, K. (2006). Ecological characteristics of small farmland ponds: Associations with land use practices at multiple spatial scales. *Biological Conservation*, *131*(4), 523–532. https://doi.org/10.1016/j.biocon.2006.02.024

DEFRA. (2022, October 5). *Natural Capital and Ecosystem Assessment Programme*. Policy Paper. Natural Capital and Ecosystem Assessment Programme

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. In *Molecular Ecology* (Vol. 26, Issue 21, pp. 5872–5895). Blackwell Publishing Ltd. https://doi.org/10.1111/mec.14350

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

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

Dembowska, E. A., Mieszczankin, T., & Napiórkowski, P. (2018). Changes of the phytoplankton community as symptoms of deterioration of water quality in a shallow lake. *Environmental Monitoring and Assessment*, *190*(2), 95. https://doi.org/10.1007/s10661-018-6465-1

Dengler, J. (2009). Which function describes the species–area relationship best? A review and empirical evaluation. *Journal of Biogeography*, *36*(4), 728–744. https://doi.org/10.1111/j.1365-2699.2008.02038.x

Di Muri, C., Lawson Handley, L., Bean, C. W., Li, J., Peirson, G., Sellers, G. S., Walsh, K., Watson, H. V., Winfield, I. J., & Hänfling, B. (2020). Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds. *Metabarcoding and Metagenomics*, *4*. https://doi.org/10.3897/mbmg.4.56959

Diamond, J. M. (1975). The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, *7*(2), 129–146. https://doi.org/10.1016/0006-3207(75)90052-X

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., Holdaway, R. J., Lear, G., Makiola, A., Morales, S. E., Powell, J. R., & Weaver, L. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. In *Molecular Ecology Resources* (Vol. 18, Issue 5, pp. 940–952). Blackwell Publishing Ltd. https://doi.org/10.1111/1755-0998.12907

Didham, R. K., Blakely, T. J., Ewers, R. M., Hitchings, T. R., Ward, J. B., & Winterbourn, M. J. (2012). Horizontal and vertical structuring in the dispersal of adult aquatic insects in a fragmented landscape. *Fundamental and Applied Limnology*, *180*(1), 27–40. https://doi.org/10.1127/1863-9135/2012/0243

Doherty, M., Yager, P. L., Moran, M. A., Coles, V. J., Fortunato, C. S., Krusche, A. V., Medeiros, P. M., Payet, J. P., Richey, J. E., Satinsky, B. M., Sawakuchi, H. O., Ward, N. D., & Crump, B. C.

(2017). Bacterial Biogeography across the Amazon River-Ocean Continuum. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.00882

Doi, H., Akamatsu, Y., Goto, M., Inui, R., Komuro, T., Nagano, M., & Minamoto, T. (2021). Broad-scale detection of environmental DNA for an invasive macrophyte and the relationship between DNA concentration and coverage in rivers. *Biological Invasions*, *23*(2), 507–520. https://doi.org/10.1007/s10530-020-02380-9

Doi, H., Takahara, T., Minamoto, T., Matsuhashi, S., Uchii, K., & Yamanaka, H. (2015). Droplet Digital Polymerase Chain Reaction (PCR) Outperforms Real-Time PCR in the Detection of Environmental DNA from an Invasive Fish Species. *Environmental Science & Technology*, *49*(9), 5601–5608. https://doi.org/10.1021/acs.est.5b00253

Dormontt, E. E., van Dijk, K. J., Bell, K. L., Biffin, E., Breed, M. F., Byrne, M., Caddy-Retalic, S., Encinas-Viso, F., Nevill, P. G., Shapcott, A., Young, J. M., Waycott, M., & Lowe, A. J. (2018). Advancing DNA barcoding and metabarcoding applications for plants requires systematic analysis of herbarium collections-an Australian perspective. *Frontiers in Ecology and Evolution*, *6*(SEP), 134. https://doi.org/10.3389/fevo.2018.00134

Downing, J. A. (2010). Emerging global role of small lakes and ponds: little things mean a lot. *Limnetica*, *1*(29). http://revistes.iec.cat/index.php/limnetica/article/view/39973

Downing, J. A., Cole, J. J., Middelburg, J. J., Striegl, R. G., Duarte, C. M., Kortelainen, P., Prairie, Y. T., & Laube, K. A. (2008). Sediment organic carbon burial in agriculturally eutrophic impoundments over the last century. *Global Biogeochemical Cycles*, *22*(1). https://doi.org/10.1029/2006GB002854

Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., McDowell, W. H., Kortelainen, P., Caraco, N. F., Melack, J. M., & Middelburg, J. J. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, *51*(5), 2388–2397. https://doi.org/10.4319/lo.2006.51.5.2388

Drummond, J. A., Larson, E. R., Li, Y., Lodge, D. M., Gantz, C. A., Pfrender, M. E., Renshaw, M. A., Correa, A. M. S., & Egan, S. P. (2021). Diversity Metrics Are Robust to Differences in Sampling Location and Depth for Environmental DNA of Plants in Small Temperate Lakes. *Frontiers in Environmental Science*, *9*. https://doi.org/10.3389/fenvs.2021.617924

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*(02), 163. https://doi.org/10.1017/S1464793105006950

Dunker, K. J., Sepulveda, A. J., Massengill, R. L., Olsen, J. B., Russ, O. L., Wenburg, J. K., & Antonovich, A. (2016). Potential of Environmental DNA to Evaluate Northern Pike (Esox lucius) Eradication Efforts: An Experimental Test and Case Study. *PLOS ONE*, *11*(9), e0162277. https://doi.org/10.1371/journal.pone.0162277

Eiler, A., Löfgren, A., Hjerne, O., Nordén, S., & Saetre, P. (2018). Environmental DNA (eDNA) detects the pool frog (Pelophylax lessonae) at times when traditional monitoring methods are insensitive. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-23740-5

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*(7), e0130324. https://doi.org/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*. https://doi.org/10.3389/fenvs.2017.00011

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

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275. https://doi.org/10.1111/2041-210X.12789

Environment Agency, E. (2021, November 18). *Ecological status of surface waters in Europe*.

Eriksen, T. E., Brittain, J. E., Søli, G., Jacobsen, D., Goethals, P., & Friberg, N. (2021). A global perspective on the application of riverine macroinvertebrates as biological indicators in Africa, South-Central America, Mexico and Southern Asia. *Ecological Indicators*, *126*, 107609. https://doi.org/10.1016/j.ecolind.2021.107609

European Environment Agency. (2021, November 18). *Ecological Status of Surface Waters in Europe*.

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., & Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. In T. Poisot (Ed.), *Functional Ecology* (Vol. 30, Issue 12, pp. 1904–1916). Blackwell Publishing Ltd. https://doi.org/10.1111/1365-2435.12659

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., Jerde, C. L., Lodge, D. M., Lamberti, G. A., & Pfrender, M. E. (2017). Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Science*. https://doi.org/10.1139/cjfas-2016-0306

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-Scale Monitoring of Plants through Environmental DNA Metabarcoding of Soil: Recovery, Resolution, and Annotation of Four DNA Markers. *PLOS ONE*, *11*(6), e0157505. https://doi.org/10.1371/journal.pone.0157505

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 487–515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

Fahrig, L. (2013). Rethinking patch size and isolation effects: the habitat amount hypothesis. *Journal of Biogeography*, *40*(9), 1649–1663. https://doi.org/10.1111/jbi.12130

Fahrig, L. (2017). Ecological Responses to Habitat Fragmentation Per Se. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 1–23. https://doi.org/10.1146/annurev-ecolsys-110316-022612

Fahrig, L. (2020). Why do several small patches hold more species than few large patches? *Global Ecology and Biogeography*, *29*(4), 615–628. https://doi.org/10.1111/geb.13059

Fediajevaite, J., Priestley, V., Arnold, R., & Savolainen, V. (2021). Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. *Ecology and Evolution*, *11*(9), 4803–4815. https://doi.org/10.1002/ece3.7382

Fernández, S., Rodríguez, S., Martínez, J. L., Borrell, Y. J., Ardura, A., & García-Vázquez, E. (2018). Evaluating freshwater macroinvertebrates from eDNA metabarcoding: A river Nalón case study. *PLOS ONE*, *13*(8), e0201741. https://doi.org/10.1371/journal.pone.0201741

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

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

Finlayson, M., Davidson, N., Fennessy, S., Coates, D., Gardner Contributing Authors, R. C., Darwall, W., Dema, M., Everard, M., McRae, L., Perennou, C., Stroud, D., Bambaradeniya, C., Davies, P., Fang, W.-T., Hilomen, V., Kulindwa, K., Martinez, L., Ricaurte, L., Scoullos, M., … Dudley, N. (2018). *Global Wetland Outlook: State of the World's Wetlands and their Services to People*.

Fitter, R. S. Richmond., & Manuel, R. L. (1986). *Field guide to the freshwater life of Britain and north-west Europe*. Collins.

Fonseca, A., Espinoza, C., Nielsen, L. P., Marshall, I. P. G., & Gallardo, V. A. (2022). Bacterial community of sediments under the Eastern Boundary Current System shows high microdiversity and a latitudinal spatial pattern. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.1016418

Francioli, D., Lentendu, G., Lewin, S., & Kolb, S. (2021). DNA metabarcoding for the characterization of terrestrial microbiota—pitfalls and solutions. In *Microorganisms* (Vol. 9, Issue 2, pp. 1–29). MDPI AG. https://doi.org/10.3390/microorganisms9020361

Freshwater Habitats Trust. (2015). *Pond Habitat Survey Method Booklet*.

Freshwater Habitats Trust. (2019). *Pinkhill Meadow - what are the ponds like after 29 years?* https://freshwaterhabitats.org.uk/pinkhill-meadow-after-29-years/

Frühe, L., Cordier, T., Dully, V., Breiner, H., Lentendu, G., Pawlowski, J., Martins, C., Wilding, T. A., & Stoeck, T. (2021). Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. *Molecular Ecology*, *30*(13), 2988–3006. https://doi.org/10.1111/mec.15434

FUENTES-RODRÍGUEZ, F., JUAN, M., GALLEGO, I., LUSI, M., FENOY, E., LEÓN, D., PEÑALVER, P., TOJA, J., & CASAS, J. J. (2013). Diversity in Mediterranean farm ponds: trade-offs and synergies between irrigation modernisation and biodiversity conservation. *Freshwater Biology*, *58*(1), 63–78. https://doi.org/10.1111/fwb.12038

Fujiwara, A., Matsuhashi, S., Doi, H., Yamamoto, S., & Minamoto, T. (2016). *Use of environmental DNA to survey the distribution of an invasive submerged plant in ponds*. https://doi.org/10.1086/685882

Gallego, I., Davidson, T. A., Jeppesen, E., Pérez-Martínez, C., Fuentes-Rodríguez, F., Juan, M., & Casas, J. J. (2014). Disturbance from pond management obscures local and regional drivers of assemblages of primary producers. *Freshwater Biology*, *59*(7), 1406–1422. https://doi.org/10.1111/fwb.12354

Gan, H. M., Hudson, A. O., Rahman, A. Y. A., Chan, K. G., & Savka, M. A. (2013). Comparative genomic analysis of six bacteria belonging to the genus Novosphingobium: insights into marine adaptation, cell-cell signaling and bioremediation. *BMC Genomics*, *14*(1), 431. https://doi.org/10.1186/1471-2164-14-431

Gao, Y., & Wu, M. (2023). Accounting for 16S rRNA copy number prediction uncertainty and its implications in bacterial diversity analyses. *ISME Communications*, *3*(1), 59. https://doi.org/10.1038/s43705-023-00266-0

García-Girón, J., Wilkes, M., Fernández-Aláez, M., & Fernández-Aláez, C. (2019). Processes structuring macrophyte metacommunities in Mediterranean ponds: Combining novel methods to disentangle the role of dispersal limitation, species sorting and spatial scales. *Journal of Biogeography*, *46*(3), 646–656. https://doi.org/10.1111/jbi.13516

García-Machado, E., Normandeau, E., Côté, G., & Bernatchez, L. (2023). How <scp>eDNA</scp> data filtration, sequence coverage, and primer selection influence assessment of fish communities in northern temperate lakes. *Environmental DNA*. https://doi.org/10.1002/edn3.444

Garrido-Sanz, L., Senar, M. À., & Piñol, J. (2022). Relative species abundance estimation in artificial mixtures of insects using mito-metagenomics and a correction factor for the mitochondrial DNA copy number. *Molecular Ecology Resources*, *22*(1), 153–167. https://doi.org/10.1111/1755-0998.13464

Gibbons, S. M., Jones, E., Bearquiver, A., Blackwolf, F., Roundstone, W., Scott, N., Hooker, J., Madsen, R., Coleman, M. L., & Gilbert, J. A. (2014). Human and Environmental Impacts on River Sediment Microbial Communities. *PLoS ONE*, *9*(5), e97435. https://doi.org/10.1371/journal.pone.0097435

Giebner, H., Langen, K., Bourlat, S. J., Kukowka, S., Mayer, C., Astrin, J. J., Misof, B., & Fonseca, V. G. (2020). Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. *Molecular Ecology Resources*, *20*(5), 1333–1345. https://doi.org/10.1111/1755-0998.13201

Gilbert, J. A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C. T., Brown, C. T., Desai, N., Eisen, J. A., Evers, D., Field, D., Feng, W., Huson, D., Jansson, J., Knight, R., Knight, J., Kolker, E.,

Konstantindis, K., Kostka, J., Kyrpides, N., ... Stevens, R. (2010). Meeting Report: The Terabase Metagenomics Workshop and the Vision of an Earth Microbiome Project. *Standards in Genomic Sciences*, *3*(3), 243–248. https://doi.org/10.4056/sigs.1433550

Gingera, T. D., Bajno, R., Docker, M. F., & Reist, J. D. (2017). Open Access Special Issue: Management of Invasive Species in Inland Waters Environmental DNA as a detection tool for zebra mussels Dreissena polymorpha (Pallas, 1771) at the forefront of an invasion event in Lake Winnipeg, Manitoba, Canada. *Management of Biological Invasions*, *8*, 287–300. https://doi.org/10.3391/mbi.2017.8.3.03

Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., & Field, K. G. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, *345*(6270), 60–63. https://doi.org/10.1038/345060a0

Glen-Bott, J. I. (1951). Helicodendron giganteum n.sp. and other aerial-sporing Hyphomycetes of submerged dead leaves. *Transactions of the British Mycological Society*, *34*(3), 275-IN4. https://doi.org/10.1016/S0007-1536(51)80052-4

Goldberg, C. S., Strickler, K. M., & Fremier, A. K. (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. https://doi.org/10.1016/j.scitotenv.2018.02.295

Golden, H. E., Lane, C. R., Amatya, D. M., Bandilla, K. W., Raanan Kiperwas, H., Knightes, C. D., & Ssegane, H. (2014). Hydrologic connectivity between geographically isolated wetlands and surface water systems: A review of select modeling methods. In *Environmental Modelling and Software* (Vol. 53, pp. 190–206). Elsevier. https://doi.org/10.1016/j.envsoft.2013.12.004

Gomez, F. (2020). DIVERSITY AND CLASSIFICATION OF DINOFLAGELLATES. In D. V. Subba-Rao (Ed.), *Dinoflagellates : classification, evolution, physiology and ecological significance* (pp. 1–38). Nova Science.

Gong, W., & Marchetti, A. (2019). Estimation of 18S Gene Copy Number in Marine Eukaryotic Plankton Using a Next-Generation Sequencing Approach. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00219

Govindarajan, A. F., McCartin, L., Adams, A., Allan, E., Belani, A., Francolini, R., Fujii, J., Gomez-Ibañez, D., Kukulya, A., Marin, F., Tradd, K., Yoerger, D. R., McDermott, J. M., & Herrera, S. (2022). Improved biodiversity detection using a large-volume environmental DNA sampler with in situ filtration and implications for marine eDNA sampling strategies. *Deep Sea Research Part I: Oceanographic Research Papers*, *189*, 103871. https://doi.org/10.1016/j.dsr.2022.103871

Green, J., & Bohannan, B. J. M. (2006). Spatial scaling of microbial biodiversity. *Trends in Ecology & Evolution*, *21*(9), 501–507. https://doi.org/10.1016/j.tree.2006.06.012

Grimes, W. F. (1940). THE BRITISH ISLANDS AND THEIR VEGETATION. By A. G. Tansley. Cambridge University Press. 1939. pp. XXXVIII, 930, with 162 plates containing 418 photographs, and 179 text-figures . 45 s . . *Antiquity*, *14*(54), 228–231. https://doi.org/10.1017/s0003598x00015210

Grossart, H. P., Van den Wyngaert, S., Kagami, M., Wurzbacher, C., Cunliffe, M., & Rojas-Jimenez, K. (2019). Fungi in aquatic ecosystems. In *Nature Reviews Microbiology* (Vol. 17, Issue 6, pp. 339–354). Nature Publishing Group. https://doi.org/10.1038/s41579-019-0175-8

Gweon, H. S., Bowes, M. J., Moorhouse, H. L., Oliver, A. E., Bailey, M. J., Acreman, M. C., & Read, D. S. (2021). Contrasting community assembly processes structure lotic bacteria metacommunities along the river continuum. *Environmental Microbiology*, *23*(1), 484–498. https://doi.org/10.1111/1462-2920.15337

Habitats Trust, F. (2015). *Pond Habitat Survey Method Booklet*.

Haileselasie, T. H., Mergeay, J., Vanoverbeke, J., Orsini, L., & De Meester, L. (2018). Founder effects determine the genetic structure of the water flea *Daphnia* in Ethiopian reservoirs. *Limnology and Oceanography*, *63*(2), 915–926. https://doi.org/10.1002/lno.10678

Hajibabaei, M., Porter, T. M., Robinson, C. V., Baird, D. J., Shokralla, S., & Wright, M. (2019). Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. *BioRxiv*, 575928. https://doi.org/10.1101/575928

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C., & Baird, D. J. (2011). Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. *PLoS ONE*, *6*(4), e17497. https://doi.org/10.1371/journal.pone.0017497

Halley, J. M., Monokrousos, N., Mazaris, A. D., Newmark, W. D., & Vokou, D. (2016). Dynamics of extinction debt across five taxonomic groups. *Nature Communications*, *7*(1), 1–6. https://doi.org/10.1038/ncomms12283

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C., Oliver, A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, *25*(13), 3101–3119. https://doi.org/10.1111/mec.13660

Hanski, I. (1998). Metapopulation dynamics. In *Nature* (Vol. 396, Issue 6706, pp. 41–49). https://doi.org/10.1038/23876

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506. https://doi.org/10.1038/nrmicro2795

Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., Read, D. S., Watson, H. V., Sayer, C. D., Jones, E. P., Priestley, V., Mächler, E., Múrria, C., Garcés-Pastor, S., Medupin, C., Burgess, K., Benson, G., Boonham, N., Griffiths, R. A., … Hänfling, B. (2019). Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. In *Hydrobiologia* (Vol. 826, Issue 1, pp. 25–41). Springer International Publishing. https://doi.org/10.1007/s10750-018-3750-5

Harper, L. R., Handley, L., Hahn, C., Boonham, N., Rees, H. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S., Hänfling, B., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S., & Hänfling, B. (2019). Generating

and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: A case study on a threatened amphibian. *Environmental DNA*, edn3.57. https://doi.org/10.1002/edn3.57

Harper, L. R., Handley, L. L., Carpenter, A. I., Murray-Dickson, G., Muri, C. Di, Macgregor, C. J., Logan, T. W., Law, A., Breithaupt, T., Read, D. S., McDevitt, A., & Hänfling, B. (2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *BioRxiv*, 546218. https://doi.org/10.1101/546218

Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E., Adams, I. P., 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*(12), 6330–6341. https://doi.org/10.1002/ece3.4013

Harper, L. R., Lawson Handley, L., Sayer, C. D., Read, D. S., Benucci, M., Blackman, R. C., Hill, M. J., & Hänfling, B. (2021). Assessing the impact of the threatened crucian carp (*Carassius carassius*) on pond invertebrate diversity: A comparison of conventional and molecular tools. *Molecular Ecology*, *30*(13), 3252–3269. https://doi.org/10.1111/mec.15670

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1915), 20191409. https://doi.org/10.1098/rspb.2019.1409

Hassall, C., Hollinshead, J., & Hull, A. (2012). Temporal dynamics of aquatic communities and implications for pond conservation. *Biodiversity and Conservation*, *21*(3), 829–852. https://doi.org/10.1007/s10531-011-0223-9

Hawksworth, D. L., & Lücking, R. (2017). Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiology Spectrum*, *5*(4). https://doi.org/10.1128/microbiolspec.FUNK-0052-2016

Hayden, C. J., & Beman, J. M. (2016). Microbial diversity and community structure along a lake elevation gradient in <scp>Y</scp> osemite <scp>N</scp> ational <scp>P</scp> ark, <scp>C</scp> alifornia, <scp>USA</scp>. *Environmental Microbiology*, *18*(6), 1782–1791. https://doi.org/10.1111/1462-2920.12938

Heckmann, K., & Schmidt, H. J. (1987). Polynucleobacter necessarius gen. nov., sp. nov., an Obligately Endosymbiotic Bacterium Living in the Cytoplasm of Euplotes aediculatus. *International Journal of Systematic Bacteriology*, *37*(4), 456–457. https://doi.org/10.1099/00207713-37-4-456

Heino, J., Melo, A. S., Siqueira, T., Soininen, J., Valanko, S., & Bini, L. M. (2015). Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater Biology*, *60*(5), 845–869. https://doi.org/10.1111/fwb.12533

Henson, M. W., Hanssen, J., Spooner, G., Fleming, P., Pukonen, M., Stahr, F., & Thrash, J. C. (2018). Nutrient dynamics and stream order influence microbial community patterns along a 2914 kilometer transect of the Mississippi River. *Limnology and Oceanography*, *63*(5), 1837–1855. https://doi.org/10.1002/lno.10811

Hestetun, J. T., Bye-Ingebrigtsen, E., Nilsson, R. H., Glover, A. G., Johansen, P.-O., & Dahlgren, T. G. (2020). Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Marine Biodiversity*, *50*(5), 70. https://doi.org/10.1007/s12526-020-01093-5

Hill, M. J., Biggs, J., Thornhill, I., Briers, R. A., Gledhill, D. G., White, J. C., Wood, P. J., & Hassall, C. (2017a). Urban ponds as an aquatic biodiversity resource in modified landscapes. *Global Change Biology*, *23*(3), 986–999. https://doi.org/10.1111/GCB.13401

Hill, M. J., Biggs, J., Thornhill, I., Briers, R. A., Gledhill, D. G., White, J. C., Wood, P. J., & Hassall, C. (2017b). Urban ponds as an aquatic biodiversity resource in modified landscapes. *Global Change Biology*, *23*(3), 986–999. https://doi.org/10.1111/gcb.13401

Hill, M. J., Hassall, C., Oertli, B., Fahrig, L., Robson, B. J., Biggs, J., Samways, M. J., Usio, N., Takamura, N., Krishnaswamy, J., & Wood, P. J. (2018). New policy directions for global pond conservation. *Conservation Letters*, *11*(5), e12447. https://doi.org/10.1111/conl.12447

Hill, M. J., Sayer, C. D., & Wood, P. J. (2016). When is the best time to sample aquatic macroinvertebrates in ponds for biodiversity assessment? *Environmental Monitoring and Assessment*, *188*(3), 194. https://doi.org/10.1007/s10661-016-5178-6

Hill, M. J., & Wood, P. J. (2014). The macroinvertebrate biodiversity and conservation value of garden and field ponds along a rural-urban gradient. *Fundamental and Applied Limnology*, *185*(1), 107–119. https://doi.org/10.1127/fal/2014/0612

Hill, M. J., Wood, P. J., Fairchild, W., Williams, P., Nicolet, P., & Biggs, J. (2021). Garden pond diversity: Opportunities for urban freshwater conservation. *Basic and Applied Ecology*, *57*, 28–40. https://doi.org/10.1016/j.baae.2021.09.005

Hindshaw, R. S., Lindsay, M. R., & Boyd, E. S. (2017). Diversity and abundance of microbial eukaryotes in stream sediments from Svalbard. *Polar Biology*, *40*(9), 1835–1843. https://doi.org/10.1007/s00300-017-2106-3

Hinlo, R., Gleeson, D., Lintermans, M., & Furlan, E. (2017). Methods to maximise recovery of environmental DNA from water samples. *PLOS ONE*, *12*(6), e0179251. https://doi.org/10.1371/journal.pone.0179251

Holgerson, M. A., & Raymond, P. A. (2016). Large contribution to inland water CO2 and CH4 emissions from very small ponds. *Nature Geoscience*, *9*(3), 222–226. https://doi.org/10.1038/ngeo2654

Holt, E., & Miller, S. (2010). Bioindicators: Using Organisms to Measure Environmental Impacts. *Nature Education Knowledge*, *3*(10:8).

Horton, D. J., Kershner, M. W., & Blackwood, C. B. (2017). Suitability of PCR primers for characterizing invertebrate communities from soil and leaf litter targeting metazoan 18S ribosomal or cytochrome oxidase I (COI) genes. *European Journal of Soil Biology*, *80*, 43–48. https://doi.org/10.1016/j.ejsobi.2017.04.003

Horváth, Z., Ptacnik, R., Vad, C. F., & Chase, J. M. (2019). Habitat loss over six decades accelerates regional and local biodiversity loss via changing landscape connectance. *Ecology Letters*, *22*(6), 1019–1027. https://doi.org/10.1111/ele.13260

Hunter, M. E., Ferrante, J. A., Meigs-Friend, G., & Ulmer, A. (2019). Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Scientific Reports*, *9*(1), 5259. https://doi.org/10.1038/s41598-019-40977-w

Hupało, K., Schmidt, S., Macher, T.-H., Weiss, M., & Leese, F. (2022a). Fresh insights into Mediterranean biodiversity: environmental DNA reveals spatio-temporal patterns of stream invertebrate communities on Sicily. *Hydrobiologia*, *849*(1), 155–173. https://doi.org/10.1007/s10750-021-04718-3

Hupało, K., Schmidt, S., Macher, T.-H., Weiss, M., & Leese, F. (2022b). Fresh insights into Mediterranean biodiversity: environmental DNA reveals spatio-temporal patterns of stream invertebrate communities on Sicily. *Hydrobiologia*, *849*(1), 155–173. https://doi.org/10.1007/s10750-021-04718-3

Incagnone, G., Marrone, F., Barone, R., Robba, L., & Naselli-Flores, L. (2015). How do freshwater organisms cross the “dry ocean”? A review on passive dispersal and colonization processes with a special focus on temporary ponds. In *Hydrobiologia* (Vol. 750, Issue 1, pp. 103–123). Kluwer Academic Publishers. https://doi.org/10.1007/s10750-014-2110-3

Interagency Freshwater Group. (2015). *Common Standards Monitoring Guidance for Freshwater Lakes*.

Ionescu, D., Bizic, M., Karnatak, R., Musseau, C. L., Onandia, G., Kasada, M., Berger, S. A., Nejstgaard, J. C., Ryo, M., Lischeid, G., Gessner, M. O., Wollrab, S., & Grossart, H. -P. (2022). From microbes to mammals: Pond biodiversity homogenization across different land-use types in an agricultural landscape. *Ecological Monographs*, *92*(3). https://doi.org/10.1002/ecm.1523

Ishige, T., Miya, M., Ushio, M., Sado, T., Ushioda, M., Maebashi, K., Yonechi, R., Lagan, P., & Matsubayashi, H. (2017). Tropical-forest mammals as detected by environmental DNA at natural saltlicks in Borneo. *Biological Conservation*, *210*, 281–285. https://doi.org/10.1016/j.biocon.2017.04.023

Ismail, A. H., & Mohd Adnan, A. A. (2016). Zooplankton Composition and Abundance as Indicators of Eutrophication in Two Small Man-made Lakes. *Tropical Life Sciences Research*, *27*(3), 31–38. https://doi.org/10.21315/tlsr2016.27.3.5

Isobe, K., Allison, S. D., Khalili, B., Martiny, A. C., & Martiny, J. B. H. (2019). Phylogenetic conservation of bacterial responses to soil nitrogen addition across continents. *Nature Communications*, *10*(1), 2499. https://doi.org/10.1038/s41467-019-10390-y

Ittner, L. D., Junghans, M., & Werner, I. (2018). Aquatic Fungi: A Disregarded Trophic Level in Ecological Risk Assessment of Organic Fungicides. *Frontiers in Environmental Science*, *6*. https://doi.org/10.3389/fenvs.2018.00105

Jenkins, D. G., Brescacin, C. R., Duxbury, C. V., Elliott, J. A., Evans, J. A., Grablow, K. R., Hillegass, M., Lyon, B. N., Metzger, G. A., Olandese, M. L., Pepe, D., Silvers, G. A., Suresch,

H. N., Thompson, T. N., Trexler, C. M., Williams, G. E., Williams, N. C., & Williams, S. E. (2007). Does size matter for dispersal distance? *Global Ecology and Biogeography*, *16*(4), 415–425. https://doi.org/10.1111/j.1466-8238.2007.00312.x

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150–157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

JGI Mycocosm. (n.d.). *Stagonospora sp. SRC1lsM3a v1.0*. 2023.

Ji, F., Yan, L., Yan, S., Qin, T., Shen, J., & Zha, J. (2021). Estimating aquatic plant diversity and distribution in rivers from Jingjinji region, China, using environmental DNA metabarcoding and a traditional survey method. *Environmental Research*, *199*, 111348. https://doi.org/10.1016/j.envres.2021.111348

Jiao, C., Zhao, D., Huang, R., He, F., & Yu, Z. (2021). Habitats and seasons differentiate the assembly of bacterial communities along a trophic gradient of freshwater lakes. *Freshwater Biology*, *66*(8), 1515–1529. https://doi.org/10.1111/fwb.13735

JNCC. (2010). *EU Water Framework Directive*. http://jncc.defra.gov.uk/page-1375

JNCC. (2011). *UK Biodiversity Action Plan - Priority Habitat Descriptions (Updated December 2011)*. http://www.jncc.gov.uk/page-5155

JNCC - Adviser to Government on Nature Conservation. (2020). *UKBI - B7. Surface water status*. https://jncc.gov.uk/our-work/ukbi-b7-surface-water-status/

Jo, T., Fukuoka, A., Uchida, K., Ushimaru, A., & Minamoto, T. (2020). Multiplex real-time PCR enables the simultaneous detection of environmental DNA from freshwater fishes: a case study of three exotic and three threatened native fishes in Japan. *Biological Invasions*, *22*(2), 455–471. https://doi.org/10.1007/s10530-019-02102-w

Jones, C. N., Nelson, N. G., & Smith, L. L. (2019). Featured Collection Introduction: The Emerging Science of Aquatic System Connectivity I. *JAWRA Journal of the American Water Resources Association*, *55*(2), 287–293. https://doi.org/10.1111/1752-1688.12739

Jones, J. I., Murphy, J. F., Anthony, S. G., Arnold, A., Blackburn, J. H., Duerdoth, C. P., Hawczak, A., Hughes, G. O., Pretty, J. L., Scarlett, P. M., Gooday, R. D., Zhang, Y. S., Fawcett, L. E., Simpson, D., Turner, A. W. B., Naden, P. S., & Skates, J. (2017). Do agri-environment schemes result in improved water quality? *Journal of Applied Ecology*, *54*(2), 537–546. https://doi.org/10.1111/1365-2664.12780

JØRGENSEN, T., HAILE, J., MÖLLER, P., ANDREEV, A., BOESSENKOOL, S., RASMUSSEN, M., KIENAST, F., COISSAC, E., TABERLET, P., BROCHMANN, C., BIGELOW, N. H., ANDERSEN, K., ORLANDO, L., GILBERT, M. T. P., & WILLERSLEV, E. (2012). A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology*, *21*(8), 1989–2003. https://doi.org/10.1111/j.1365-294X.2011.05287.x

Joy, M., & Death, R. (2002). Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology*, *47*, 2261–2275. https://doi.org/10.1046/j.1365-2427.2002.00954.x

Joye, D. A., Oertli, B., Lehmann, A., Juge, R., & Lachavanne, J.-B. (2006). The prediction of macrophyte species occurrence in Swiss ponds. In *Macrophytes in Aquatic Ecosystems: From Biology to Management* (pp. 175–182). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5390-0_25

Jurat-Fuentes, J. L., & Jackson, T. A. (2012). Bacterial Entomopathogens. In *Insect Pathology* (pp. 265–349). Elsevier. https://doi.org/10.1016/B978-0-12-384984-7.00008-7

Kamoroff, C., & Goldberg, C. S. (2017). Using environmental DNA for early detection of amphibian chytrid fungus Batrachochytrium dendrobatidis prior to a ranid die-off. *Diseases of Aquatic Organisms*, *127*(1), 75–79. https://doi.org/10.3354/dao03183

Kämpfer, P., Busse, H.-J., Longaric, I., Rosselló-Móra, R., Galatis, H., & Lodders, N. (2012). Pseudarcicella hirudinis gen. nov., sp. nov., isolated from the skin of the medical leech Hirudo medicinalis. *International Journal of Systematic and Evolutionary Microbiology*, *62*(Pt_9), 2247–2251. https://doi.org/10.1099/ijs.0.037390-0

Karr, J., Fausch, K., Angermeier, P., Yant, P., & Schlosser, I. (1986). Assessing biological integrity in running waters. A method and its rationale. *Ill Nat Hist Surv Spec Publ*, *5*.

Kelly, M. (2013). Data rich, information poor? Phytobenthos assessment and the Water Framework Directive. *European Journal of Phycology*, *48*(4), 437–450. https://doi.org/10.1080/09670262.2013.852694

Kelly-Quinn, M., Biggs, J. N., Brooks, S., Fortuño, P., Hegarty, S., Jones, J. I., & Regan, F. (2023). Opportunities, approaches and challenges to the engagement of citizens in filling small water body data gaps. *Hydrobiologia*, *850*(15), 3419–3439. https://doi.org/10.1007/s10750-022-04973-y

Kennedy, C. M., Oakleaf, J. R., Theobald, D. M., Baruch-Mordo, S., & Kiesecker, J. (2019). Managing the middle: A shift in conservation priorities based on the global human modification gradient. *Global Change Biology*, *25*(3), 811–826. https://doi.org/10.1111/gcb.14549

Kiran, M. T., & Tiwari, M. V. B. and A. (2016). Phycoremediation of Eutrophic Lakes Using Diatom Algae. In M. N. Rashed (Ed.), *Lake Sciences and Climate Change* (p. Ch. 6). IntechOpen. https://doi.org/10.5772/64111

Krienitz, L. (2009). Algae. *Encyclopedia of Inland Waters*, 103–113. https://doi.org/10.1016/B978-012370626-3.00132-0

Kueneman, J. G., Weiss, S., & McKenzie, V. J. (2017). Composition of Micro-eukaryotes on the Skin of the Cascades Frog (Rana cascadae) and Patterns of Correlation between Skin Microbes and Batrachochytrium dendrobatidis. *Frontiers in Microbiology*, *8*(DEC), 2350. https://doi.org/10.3389/fmicb.2017.02350

Kuhn, M., Lundin, E. J., Giesler, R., Johansson, M., & Karlsson, J. (2018). Emissions from thaw ponds largely offset the carbon sink of northern permafrost wetlands. *Scientific Reports*, *8*(1), 9535. https://doi.org/10.1038/s41598-018-27770-x

Kulaš, A., Gulin, V., Matoničkin Kepčija, R., Žutinić, P., Sertić Perić, M., Orlić, S., Kajan, K., Stoeck, T., Lentendu, G., Čanjevac, I., Martinić, I., & Gligora Udovič, M. (2021). Ciliates (Alveolata, Ciliophora) as bioindicators of environmental pressure: A karstic river case. *Ecological Indicators*, *124*, 107430. https://doi.org/10.1016/J.ECOLIND.2021.107430

Kuzmina, M. L., Braukmann, T. W. A., & Zakharov, E. V. (2018). Finding the pond through the weeds: eDNA reveals underestimated diversity of pondweeds. *Applications in Plant Sciences*, *6*(5), e01155. https://doi.org/10.1002/aps3.1155

Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology*, *53*(4), 1148–1157. https://doi.org/10.1111/1365-2664.12598

Ladau, J., & Eloe-Fadrosh, E. A. (2019). Spatial, Temporal, and Phylogenetic Scales of Microbial Ecology. *Trends in Microbiology*, *27*(8), 662–669. https://doi.org/10.1016/j.tim.2019.03.003

Ladin, Z. S., Ferrell, B., Dums, J. T., Moore, R. M., Levia, D. F., Shriver, W. G., D'Amico, V., Trammell, T. L. E., Setubal, J. C., & Wommack, K. E. (2021). Assessing the efficacy of eDNA metabarcoding for measuring microbial biodiversity within forest ecosystems. *Scientific Reports*, *11*(1), 1629. https://doi.org/10.1038/s41598-020-80602-9

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

Lambert, S. (2009). *Stoneworts: their habitats, ecological requirements and conservation*.

Law, A., Baker, A., Sayer, C., Foster, G., Gunn, I. D. M., Taylor, P., Pattison, Z., Blaikie, J., & Willby, N. J. (2019). The effectiveness of aquatic plants as surrogates for wider biodiversity in standing fresh waters. *Freshwater Biology*, *64*(9), 1664–1675. https://doi.org/10.1111/fwb.13369

Lawton, J. (2010). *Making Space for Nature: A Review of England's Wildlife Sites and Ecological Network*. https://www.researchgate.net/publication/268279426_Making_Space_for_Nature_A_Review_of_England's_Wildlife_Sites_and_Ecological_Network

Lear, G., Bellamy, J., Case, B. S., Lee, J. E., & Buckley, H. L. (2014). Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *The ISME Journal*, *8*(8), 1715–1726. https://doi.org/10.1038/ismej.2014.21

Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H. L., Buckley, T. R., Cruickshank, R., Dopheide, A., Handley, K. M., Hermans, S., Kamke, J., Lee, C. K., Macdiarmid, R., Morales, S. E., Orlovich, D. A., Smissen, R., Wood, J., & Holdaway, R. (2018). REVIEW Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples . *New Zealand Journal of Ecology*, *42*(1), 10. https://doi.org/10.20417/nzjecol.42.9

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., … Zimmermann, J. (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*, *2*, e11321. https://doi.org/10.3897/rio.2.e11321

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2020). Improved freshwater macroinvertebrate detection from eDNA through minimized non-target amplification Running title: Maximizing macroinvertebrate diversity detection for biomonitoring. *BioRxiv*, 2020.04.27.063545. https://doi.org/10.1101/2020.04.27.063545

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2021). Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*, *3*(1), 261–276. https://doi.org/10.1002/EDN3.177

Leibold, M. A. (1999). Biodiversity and nutrient enrichment in pond plankton communities. In *Evolutionary Ecology Research* (Vol. 1).

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., & Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, *7*(7), 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

León, D., Peñalver, P., Casas, J., Juan, M., Fuentes, F., Gallego, I., & Toja, J. (2008). Zooplankton richness in farm ponds of Andalusia (southern Spain). A comparison with natural wetlands. *Limnetica*, *29*(1), 153–162. https://doi.org/10.23818/limn.29.12

Leray, M., Ho, S.-L., Lin, I.-J., & Machida, R. J. (2018). MIDORI server: a webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. *Bioinformatics*, *34*(21), 3753–3754. https://doi.org/10.1093/bioinformatics/bty454

Lewis, B., Griffiths, R. A., & Wilkinson, J. W. (2017). *Population status of great crested newts (Triturus cristatus) at sites subjected to development mitigation Herpetological Journal* (Vol. 27).

Lewis-Phillips, J., Brooks, S. J., Sayer, C. D., McCrea, R., Siriwardena, G., Robson, H., Harrison, A. L., & Axmacher, J. C. (2019). Seasonal benefits of farmland pond management for birds. *Bird Study*, 1–11. https://doi.org/10.1080/00063657.2019.1688762

Lewis-Phillips, J., Brooks, S. J., Sayer, C. D., Patmore, I. R., Hilton, G. M., Harrison, A., Robson, H., & Axmacher, J. C. (2020). Ponds as insect chimneys: Restoring overgrown farmland ponds benefits birds through elevated productivity of emerging aquatic insects. *Biological Conservation*, *241*, 108253. https://doi.org/10.1016/j.biocon.2019.108253

Li, F., Altermatt, F., Yang, J., An, S., Li, A., & Zhang, X. (2020). Human activities' fingerprint on multitrophic biodiversity and ecosystem functions across a major river catchment in China. *Global Change Biology*, gcb.15357. https://doi.org/10.1111/gcb.15357

Li, H., Zeng, J., Ren, L., Wang, J., Xing, P., & Wu, Q. L. (2017). Contrasting patterns of diversity of abundant and rare bacterioplankton in freshwater lakes along an elevation gradient. *Limnology and Oceanography*, *62*(4), 1570–1585. https://doi.org/10.1002/lno.10518

Li, J., Lawson Handley, L.-J., Read, D. S., & Hänfling, B. (2018). The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. *Molecular Ecology Resources*, *18*(5), 1102–1114. https://doi.org/10.1111/1755-0998.12899

Li, L., Zheng, B., & Liu, L. (2010). Biomonitoring and Bioindicators Used for River Ecosystems: Definitions, Approaches and Trends. *Procedia Environmental Sciences*, *2*, 1510–1524. https://doi.org/10.1016/j.proenv.2010.10.164

Li, X., Li, F., Min, X., Xie, Y., & Zhang, Y. (2023). Embracing eDNA and machine learning for taxonomy-free microorganisms biomonitoring to assess the river ecological status. *Ecological Indicators*, *155*, 110948. https://doi.org/10.1016/j.ecolind.2023.110948

Li, Y., Yu, Z., Ji, S., Meng, J., Kong, Q., Wang, R., & Liu, J. (2021). Diverse drivers of phytoplankton dynamics in different phyla across the annual cycle in a freshwater lake. *Journal of Freshwater Ecology*, *36*(1), 13–29. https://doi.org/10.1080/02705060.2020.1868586

Lindberg, P., Lindblad, P., & Cournac, L. (2004). Gas Exchange in the Filamentous Cyanobacterium *Nostoc punctiforme* Strain ATCC 29133 and ItsHydrogenase-Deficient Mutant StrainNHM5. *Applied and Environmental Microbiology*, *70*(4), 2137–2145. https://doi.org/10.1128/AEM.70.4.2137-2145.2004

Littlefair, J. E., Allerton, J. J., Brown, A. S., Butterfield, D. M., Robins, C., Economou, C. K., Garrett, N. R., & Clare, E. L. (2023). Air-quality networks collect environmental DNA with the potential to measure biodiversity at continental scales. *Current Biology*, *33*(11), R426–R428. https://doi.org/10.1016/j.cub.2023.04.036

Liu, B., & Stevenson, R. J. (2017). Improving assessment accuracy for lake biological condition by classifying lakes with diatom typology, varying metrics and modeling multimetric indices. *Science of The Total Environment*, *609*, 263–271. https://doi.org/10.1016/j.scitotenv.2017.07.152

Liu, B., Wu, J., Hu, Y., Wang, G., & Chen, Y. (2022). Seven Years Study of the Seasonal Dynamics of Zooplankton Communities in a Large Subtropical Floodplain Ecosystem: A Test of the PEG Model. *International Journal of Environmental Research and Public Health*, *19*(2), 956. https://doi.org/10.3390/ijerph19020956

Liu, C., Cui, Y., Li, X., & Yao, M. (2021). *microeco* : an R package for data mining in microbial community ecology. *FEMS Microbiology Ecology*, *97*(2). https://doi.org/10.1093/femsec/fiaa255

Liu, K., Hu, A., Hou, L., Zhang, L., Zeng, Q., Yan, Q., Wang, F., Zhang, Z., Xiao, X., Junzhi, L., & Liu, Y. (2022). Effects of land-use patterns on the biogeography of the sediment bacteria in the Yarlung Tsangpo River. *FEMS Microbiology Ecology*, *98*(11). https://doi.org/10.1093/femsec/fiac125

Liu, Q., Zhang, H., Chang, F., Xie, P., Zhang, Y., Wu, H., Zhang, X., Peng, W., & Liu, F. (2022). eDNA revealed in situ microbial community changes in response to Trapa japonica in Lake

Qionghai and Lake Erhai, southwestern China. *Chemosphere*, *288*, 132605. https://doi.org/10.1016/j.chemosphere.2021.132605

Liu, T., Zhang, A. N., Wang, J., Liu, S., Jiang, X., Dang, C., Ma, T., Liu, S., Chen, Q., Xie, S., Zhang, T., & Ni, J. (2018). Integrated biogeography of planktonic and sedimentary bacterial communities in the Yangtze River. *Microbiome*, *6*(1), 1–14. https://doi.org/10.1186/S40168-017-0388-X/FIGURES/6

Liu, Y., Ren, Z., Qu, X., Zhang, M., Yu, Y., Zhang, Y., & Peng, W. (2020). Microbial community structure and functional properties in permanently and seasonally flooded areas in Poyang Lake. *Scientific Reports*, *10*(1), 4819. https://doi.org/10.1038/s41598-020-61569-z

Lopes, C. M., Sasso, T., Valentini, A., Dejean, T., Martins, M., Zamudio, K. R., & Haddad, C. F. B. (2017). eDNA metabarcoding: a promising method for anuran surveys in highly diverse tropical forests. *Molecular Ecology Resources*, *17*(5), 904–914. https://doi.org/10.1111/1755-0998.12643

Louca, S., Parfrey, L. W., & Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science*, *353*(6305), 1272–1277. https://doi.org/10.1126/science.aaf4507

Lovley, D. R., Ueki, T., Zhang, T., Malvankar, N. S., Shrestha, P. M., Flanagan, K. A., Aklujkar, M., Butler, J. E., Giloteaux, L., Rotaru, A.-E., Holmes, D. E., Franks, A. E., Orellana, R., Risso, C., & Nevin, K. P. (2011). *Geobacter* (pp. 1–100). https://doi.org/10.1016/B978-0-12-387661-4.00004-5

Lu, L., Zou, X., Yang, J., Xiao, Y., Wang, Y., Guo, J., & Li, Z. (2020). Biogeography of eukaryotic plankton communities along the upper Yangtze River: The potential impact of cascade dams and reservoirs. *Journal of Hydrology*, *590*, 125495. https://doi.org/10.1016/j.jhydrol.2020.125495

Luo, J., Li, X., Ma, R., Li, F., Duan, H., Hu, W., Qin, B., & Huang, W. (2016). Applying remote sensing techniques to monitoring seasonal and interannual changes of aquatic vegetation in Taihu Lake, China. *Ecological Indicators*, *60*, 503–513. https://doi.org/10.1016/J.ECOLIND.2015.07.029

Luo, M., Ji, Y., Warton, D., & Yu, D. W. (2023). Extracting abundance information from <scp>DNA</scp> -based data. *Molecular Ecology Resources*, *23*(1), 174–189. https://doi.org/10.1111/1755-0998.13703

Lv, X., Yu, J., Fu, Y., Ma, B., Qu, F., Ning, K., & Wu, H. (2014). A Meta-Analysis of the Bacterial and Archaeal Diversity Observed in Wetland Soils. *The Scientific World Journal*, *2014*, 437684. https://doi.org/10.1155/2014/437684

Lyche-Solheim, A., Feld, C. K., Birk, S., Phillips, G., Carvalho, L., Morabito, G., Mischke, U., Willby, N., Søndergaard, M., Hellsten, S., Kolada, A., Mjelde, M., Böhmer, J., Miler, O., Pusch, M. T., Argillier, C., Jeppesen, E., Lauridsen, T. L., & Poikane, S. (2013). Ecological status assessment of European lakes: a comparison of metrics for phytoplankton, macrophytes, benthic invertebrates and fish. *Hydrobiologia*, *704*(1), 57–74. https://doi.org/10.1007/s10750-012-1436-y

MACARTHUR, R. H., & WILSON, E. O. (1967). *The Theory of Island Biogeography* (REV-Revised). Princeton University Press. http://www.jstor.org/stable/j.ctt19cc1t2

Mace, G., Masundire, H., Lead, J. B., Ricketts, T., Brooks, T., Hoffmann, M., Stuart, S., Balmford, A., Purvis, A., Reyers, B., Wang, J., Revenga, C., Kennedy, E., Naeem, S., Alkemade, R., Allnutt, T., Bakarr, M., Bond, W., Chanson, J., … Williams, P. (2005). *Biodiversity in Ecosystems and Human Well-being: Current State and Trends*.

Mächler, E., Walser, J., & Altermatt, F. (2021). Decision-making and best practices for taxonomy-free environmental DNA metabarcoding in biomonitoring using Hill numbers. *Molecular Ecology*, *30*(13), 3326–3339. https://doi.org/10.1111/mec.15725

Macingo, S. C., Kormas, K. A., Oikonomou, A., & Karayanni, H. (2019). Taxa–area and distance–decay relationships of unicellular eukaryotes along an elevation gradient of mountainous freshwater ecosystems. *Journal of Plankton Research*, *41*(6), 821–834. https://doi.org/10.1093/plankt/fbz066

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). ESTIMATING SITE OCCUPANCY, COLONIZATION, AND LOCAL EXTINCTION WHEN A SPECIES IS DETECTED IMPERFECTLY. *Ecology*, *84*(8), 2200–2207. https://doi.org/10.1890/02-3090

Maguire, B. (1963). The Passive Dispersal of Small Aquatic Organisms and Their Colonization of Isolated Bodies. In *Source: Ecological Monographs* (Vol. 33, Issue 2).

Mainstone, C., Edwards, F., Hall, R., Scarlett, P., Carvalho, L., Webb, G., Taylor, P., & Laize, C. (2018). *Developing a coherent framework for assessing priority freshwater habitats in England* (Issue April).

Mainstone, C. H. (2016). *A narrative for conserving freshwater and wetland habitats in England*.

Mangot, J.-F., Domaizon, I., Taib, N., Marouni, N., Duffaud, E., Bronner, G., & Debroas, D. (2013). Short-term dynamics of diversity patterns: evidence of continual reassembly within lacustrine small eukaryotes. *Environmental Microbiology*, *15*(6), 1745–1758. https://doi.org/10.1111/1462-2920.12065

Manning, F. S., Curtis, P. J., Walker, I. R., & Pither, J. (2021). Potential long-distance dispersal of freshwater diatoms adhering to waterfowl plumage. *Freshwater Biology*, *66*(6), 1136–1148. https://doi.org/10.1111/fwb.13706

Marshall, N. T., & Stepien, C. A. (2019). Invasion genetics from eDNA and thousands of larvae: A targeted metabarcoding assay that distinguishes species and population variation of zebra and quagga mussels. *Ecology and Evolution*, *9*(6), 3515–3538. https://doi.org/10.1002/ece3.4985

Marston, C., Rowland, C. S., O'Neill, A. W., & Morton, R. D. (2022). *Land Cover Map 2021 (10 m classified pixels, GB)*. NERC EDS Environmental Information Data Centre.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Martínez-Santos, M., Lanzén, A., Unda-Calvo, J., Martín, I., Garbisu, C., & Ruiz-Romera, E. (2018). Treated and untreated wastewater effluents alter river sediment bacterial

communities involved in nitrogen and sulphur cycling. *The Science of the Total Environment*, *633*, 1051–1061. https://doi.org/10.1016/J.SCITOTENV.2018.03.229

Martínez-Sanz, C., Cenzano, C. S. S., Fernández-Aláez, M., & García-Criado, F. (2012). Relative contribution of small mountain ponds to regional richness of littoral macroinvertebrates and the implications for conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *22*(2), 155–164. https://doi.org/10.1002/aqc.2227

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*(6), e0156217. https://doi.org/10.1371/journal.pone.0156217

Matsuoka, S., Sugiyama, Y., Shimono, Y., Ushio, M., & Doi, H. (2021). Evaluation of seasonal dynamics of fungal DNA assemblages in a flow-regulated stream in a restored forest using eDNA metabarcoding. *BioRxiv*, 2020.12.10.420661. https://doi.org/10.1101/2020.12.10.420661

Matthews, J. (2010). Anthropogenic climate change impacts on ponds: a thermal mass perspective. *BioRisk*, *5*, 193–209. https://doi.org/10.3897/biorisk.5.849

Mauvisseau, Q., Coignet, A., Delaunay, C., Pinet, F., Bouchon, D., & Souty-Grosset, C. (2018). Environmental DNA as an efficient tool for detecting invasive crayfishes in freshwater ponds. *Hydrobiologia*, *805*(1), 163–175. https://doi.org/10.1007/S10750-017-3288-Y/TABLES/5

McGee, K. M., Robinson, C. V., & Hajibabaei, M. (2019). Gaps in DNA-Based Biomonitoring Across the Globe. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00337

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, *8*(4), e61217. https://doi.org/10.1371/JOURNAL.PONE.0061217

Mejbel, H. S., Dodsworth, W., & Pick, F. R. (2022). Effects of temperature and oxygen on cyanobacterial DNA preservation in sediments: A comparison study of major taxa. *Environmental DNA*, *4*(4), 717–731. https://doi.org/10.1002/edn3.289

Miaud, C., Arnal, V., Poulain, M., Valentini, A., & Dejean, T. (2019). eDNA Increases the Detectability of Ranavirus Infection in an Alpine Amphibian Population. *Viruses*, *11*(6), 526. https://doi.org/10.3390/v11060526

Mikhailov, I. S., Galachyants, Y. P., Bukin, Y. S., Petrova, D. P., Bashenkhaeva, M. V., Sakirko, M. V., Blinov, V. V., Titova, L. A., Zakharova, Y. R., & Likhoshway, Y. V. (2022). Seasonal Succession and Coherence Among Bacteria and Microeukaryotes in Lake Baikal. *Microbial Ecology*, *84*(2), 404–422. https://doi.org/10.1007/s00248-021-01860-2

Mikhailov, I. S., Zakharova, Y. R., Bukin, Y. S., Galachyants, Y. P., Petrova, D. P., Sakirko, M. V., & Likhoshway, Y. V. (2019). Co-occurrence Networks Among Bacteria and Microbial Eukaryotes of Lake Baikal During a Spring Phytoplankton Bloom. *Microbial Ecology*, *77*(1), 96–109. https://doi.org/10.1007/s00248-018-1212-2

Mitsi, K., Richter, D. J., Arroyo, A. S., López-Escardó, D., Antó, M., Oterino, A. G., & Ruiz-Trillo, I. (2023). Taxonomic composition, community structure and molecular novelty of microeukaryotes in a temperate oligomesotrophic lake as revealed by metabarcoding. *Scientific Reports*, *13*(1), 3119. https://doi.org/10.1038/s41598-023-30228-4

Moss, W. E., Harper, L. R., Davis, M. A., Goldberg, C. S., Smith, M. M., & Johnson, P. T. J. (2022). Navigating the trade-offs between environmental DNA and conventional field surveys for improved amphibian monitoring. *Ecosphere*, *13*(2), e3941. https://doi.org/10.1002/ECS2.3941

Murali, A., Bhargava, A., & Wright, E. S. (2018). IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome*, *6*(1), 140. https://doi.org/10.1186/s40168-018-0521-5

Mushet, D. M., Alexander, L. C., Bennett, M., Schofield, K., Christensen, J. R., Ali, G., Pollard, A., Fritz, K., & Lang, M. W. (2019). Differing Modes of Biotic Connectivity within Freshwater Ecosystem Mosaics. *Journal of the American Water Resources Association*, *55*(2), 307–317. https://doi.org/10.1111/1752-1688.12683

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., & Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish using <scp>eDNA</scp> metabarcoding and conventional observation methods. *Freshwater Biology*, *63*(6), 569–580. https://doi.org/10.1111/fwb.13094

Nakatsu, C. H., Byappanahalli, M. N., & Nevers, M. B. (2019). Bacterial Community 16S rRNA Gene Sequencing Characterizes Riverine Microbial Impact on Lake Michigan. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.00996

Nathan, R., & Shohami, D. (2013). Dispersal - Ecology. In D. Gibson (Ed.), *Oxford Bibliographies*. https://www.oxfordbibliographies.com/view/document/obo-9780199830060/obo-9780199830060-0033.xml

Natural England. (2019). *Natural England's Geoportal: England-wide data for great crested newts now available - Natural England*. https://naturalengland.blog.gov.uk/2019/12/17/natural-englands-geoportal-england-wide-data-for-great-crested-newts-now-available/

Natural England. (2020). *Great Crested Newts eDNA Pond Surveys for District Level Licensing (England) | Natural England Open Data Geoportal*. https://naturalengland-defra.opendata.arcgis.com/datasets/ffba3805a4d9439c95351ef7f26ab33c_0

Natural History Museum, CEFAS, ADAS, & Natural England. (2022). *GenePools 2021:Report*. https://www.waterlowpark.org.uk/wp-content/uploads/2023/02/7-GenePools-Results-Report.pdf

NatureMetrics. (2022). *Advice note SAMPLING STRATEGIES FOR AQUATIC ENVIRONMENTS*. NatureMetrics Protocols.

Ndatimana, G., Nantege, D., & Arimoro, F. O. (2023). A review of the application of the macroinvertebrate-based multimetric indices (MMIs) for water quality monitoring in lakes.

Environmental Science and Pollution Research, *30*(29), 73098–73115. https://doi.org/10.1007/s11356-023-27559-0

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356. https://doi.org/10.1128/mmbr.00051-12

Newton, J., Sepulveda, A., Sylvester, K., & Thum, R. A. (2016). Potential utility of environmental DNA for early detection of Eurasian watermilfoil (Myriophyllum spicatum). *J. Aquat. Plant Manage*, *54*, 46–49.

Nichols, S. J., Kefford, B. J., Campbell, C. D., Bylemans, J., Chandler, E., Bray, J. P., Shackleton, M., Robinson, K. L., Carew, M. E., & Furlan, E. M. (2019). Towards routine DNA metabarcoding of macroinvertebrates using bulk samples for freshwater bioassessment: Effects of debris and storage conditions on the recovery of target taxa. *Freshwater Biology*, fwb.13443. https://doi.org/10.1111/fwb.13443

Niggebrugge, K., Durance, I., Watson, A. M., Leuven, R. S. E. W., & Ormerod, S. J. (2007). Applying landscape ecology to conservation biology: Spatially explicit analysis reveals dispersal limits on threatened wetland gastropods. *Biological Conservation*, *139*(3–4), 286–296. https://doi.org/10.1016/j.biocon.2007.07.003

Nijman, T. P. A., Davidson, T. A., Weideveld, S. T. J., Audet, J., Esposito, C., Levi, E. E., Ho, A., Lamers, L. P. M., Jeppesen, E., & Veraart, A. J. (2021). Warming and eutrophication interactively drive changes in the methane-oxidizing community of shallow lakes. *ISME Communications*, *1*(1), 32. https://doi.org/10.1038/s43705-021-00026-y

Nnadozie, C. F., & Odume, O. N. (2019). Freshwater environments as reservoirs of antibiotic resistant bacteria and their role in the dissemination of antibiotic resistance genes. *Environmental Pollution*, *254*, 113067. https://doi.org/10.1016/j.envpol.2019.113067

Nuy, J. K., Hoetzinger, M., Hahn, M. W., Beisser, D., & Boenigk, J. (2020). Ecological Differentiation in Two Major Freshwater Bacterial Taxa Along Environmental Gradients. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.00154

Oertli, B., Céréghino, R., Biggs, J., Declerck, S., Hull, A., & Miracle, R. M. (2010). Pond Conservation in Europe. In *Hydrobiologia*. Springer Netherlands. https://doi.org/10.1007/978-90-481-9088-1

Oertli, B., Joye, D. A., Castella, E., Juge, R., Cambin, D., & Lachavanne, J. B. (2002). Does size matter? The relationship between pond area and biodiversity. *Biological Conservation*, *104*(1), 59–70. https://doi.org/10.1016/S0006-3207(01)00154-9

Ogram, A., Sayler, G. S., & Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, *7*(2–3), 57–66. https://doi.org/10.1016/0167-7012(87)90025-X

Oksanen, J., Simpson, G., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., hara, R., Solymos, P., STEVENS, H., Szöcs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D.,

Carvalho, G., Chirico, M., De Cáceres, M., Durand, S., & Weedon, J. (2022). *vegan community ecology package version 2.6-2 April 2022*.

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., Deiner, K., Mahon, A. R., Brueseke, M. A., Shirey, P. D., Pfrender, M. E., Lodge, D. M., & Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, *6*(12), 4214–4226. https://doi.org/10.1002/ece3.2186

Oliver, A. E., Newbold, L. K., Gweon, H. S., Read, D. S., Woodcock, B. A., & Pywell, R. F. (2021). Integration of DNA extraction, metabarcoding and an informatics pipeline to underpin a national citizen science honey monitoring scheme. *MethodsX*, *8*, 101303. https://doi.org/10.1016/J.MEX.2021.101303

O'Malley, M. A. (2008). 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *39*(3), 314–325. https://doi.org/10.1016/j.shpsc.2008.06.005

Orland, C., Yakimovich, K. M., Mykytczuk, N. C. S., Basiliko, N., & Tanentzap, A. J. (2020). Think global, act local: The small-scale environment mainly influences microbial community development and function in lake sediment. *Limnology and Oceanography*, *65*(S1). https://doi.org/10.1002/lno.11370

Ormerod, S. J., Durance, I., Hatton-Ellis, T. W., & Cable, J. (2011). *Landscape Connectivity of Freshwater Ecosystems: Strategic Review and Recommendations*. https://www.researchgate.net/publication/256662076_Landscape_Connectivity_of_Freshwater_Ecosystems_Strategic_Review_and_Recommendations

Oshkin, I. Y., Belova, S. E., Danilova, O. V., Miroshnikov, K. K., Rijpstra, W. I. C., Sinninghe Damsté, J. S., Liesack, W., & Dedysh, S. N. (2016). Methylovulum psychrotolerans sp. nov., a cold-adapted methanotroph from low-temperature terrestrial environments, and emended description of the genus Methylovulum. *International Journal of Systematic and Evolutionary Microbiology*, *66*(6), 2417–2423. https://doi.org/10.1099/ijsem.0.001046

Oswald, K., Graf, J. S., Littmann, S., Tienken, D., Brand, A., Wehrli, B., Albertsen, M., Daims, H., Wagner, M., Kuypers, M. M., Schubert, C. J., & Milucka, J. (2017). Crenothrix are major methane consumers in stratified lakes. *The ISME Journal*, *11*(9), 2124–2140. https://doi.org/10.1038/ismej.2017.77

Ovenden, D., & Greenhalgh, M. (2007). *Freshwater Life*. Collins.

Ozinga, W. A., Rö mermann, C., Bekker, R. M., Prinzing, A., M Tamis, W. L., J Schaminé, J. H., Hennekens, S. M., Thompson, K., Poschlod, P., Kleyer, M., Bakker, J. P., & van Groenendael, J. M. (n.d.). *Dispersal failure contributes to plant losses in NW Europe*. https://doi.org/10.1111/j.1461-0248.2008.01261.x

Padgett-Stewart, T. M., Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2016). An eDNA assay for river otter detection: a tool for surveying a semi-aquatic mammal. *Conservation Genetics Resources*, *8*(1), 5–7. https://doi.org/10.1007/s12686-015-0511-x

Padial, A. A., Ceschin, F., Declerck, S. A. J., De Meester, L., Bonecker, C. C., Lansac-Tôha, F. A., Rodrigues, L., Rodrigues, L. C., Train, S., Velho, L. F. M., & Bini, L. M. (2014). Dispersal Ability Determines the Role of Environmental, Spatial and Temporal Drivers of Metacommunity Structure. *PLoS ONE*, *9*(10), e111227. https://doi.org/10.1371/journal.pone.0111227

Parducci, L., Nota, K., & Wood, J. (2018). *Reconstructing Past Vegetation Communities Using Ancient DNA from Lake Sediments* (pp. 163–187). Springer, Cham. https://doi.org/10.1007/13836_2018_38

Pascault, N., Roux, S., Artigas, J., Pesce, S., Leloup, J., Tadonleke, R. D., Debroas, D., Bouchez, A., & Humbert, J.-F. (2014). A high-throughput sequencing ecotoxicology study of freshwater bacterial communities and their responses to tebuconazole. *FEMS Microbiology Ecology*, *90*(3), 563–574. https://doi.org/10.1111/1574-6941.12416

Pasztaleniec, A. (2016). An advanced phytoplankton trophic index: Test and validation with a nationwide lake survey in Poland. *International Review of Hydrobiology*, *101*(1–2), 20–35. https://doi.org/10.1002/iroh.201501799

PATEL, G. B., SPROTT, G. D., & FEIN, J. E. (1990). Isolation and Characterization of Methanobacterium espanolae sp. nov., a Mesophilic, Moderately Acidiphilic Methanogen. *International Journal of Systematic Bacteriology*, *40*(1), 12–18. https://doi.org/10.1099/00207713-40-1-12

Patriarca, A., Vaamonde, G., & Pinto, V. F. (2014). Alternaria. In *Encyclopedia of Food Microbiology* (pp. 54–60). Elsevier. https://doi.org/10.1016/B978-0-12-384730-0.00007-0

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Feio, M. J., Filipe, A. F., Fornaroli, R., Graf, W., Herder, J., van der Hoorn, B., Iwan Jones, J., Sagova-Mareckova, M., Moritz, C., … Kahlert, M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. In *Science of the Total Environment* (Vols. 637–638, pp. 1295–1310). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2018.05.002

Pearman, J. K., Adamson, J., Thomson-Laing, G., Thompson, L., Waters, S., Vandergoes, M. J., Howarth, J. D., & Wood, S. A. (2023a). Deterministic processes drive national-scale patterns in lake surface sediment bacteria and eukaryotic assemblage composition. *Limnology and Oceanography*, *68*(1), 40–55. https://doi.org/10.1002/lno.12247

Pearman, J. K., Adamson, J., Thomson-Laing, G., Thompson, L., Waters, S., Vandergoes, M. J., Howarth, J. D., & Wood, S. A. (2023b). Deterministic processes drive national-scale patterns in lake surface sediment bacteria and eukaryotic assemblage composition. *Limnology and Oceanography*, *68*(1), 40–55. https://doi.org/10.1002/lno.12247

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439. https://doi.org/10.32614/RJ-2018-009

Pfeifer, M., Lefebvre, V., Peres, C. A., Banks-Leite, C., Wearn, O. R., Marsh, C. J., Butchart, S. H. M., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley,

A., Harris, S. M., Klingbeil, B. T., Kormann, U., Lens, L., Medina-Rangel, G. F., … Ewers, R. M. (2017). Creation of forest edges has a global impact on forest vertebrates. *Nature*, *551*(7679), 187–191. https://doi.org/10.1038/nature24457

Phalan, B., Onial, M., Balmford, A., & Green, R. E. (2011). Reconciling Food Production and Biodiversity Conservation: Land Sharing and Land Sparing Compared. *Science*, *333*(6047), 1289–1291. https://doi.org/10.1126/science.1208742

Phillips, G., Willby, N., & Moss, B. (2016). Submerged macrophyte decline in shallow lakes: What have we learnt in the last forty years? *Aquatic Botany*, *135*, 37–45. https://doi.org/10.1016/j.aquabot.2016.04.004

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Humphrey, J. S., Keacher, K. L., Bruce, W. E., & Avery, M. L. (2014). Detecting an elusive invasive species: A diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources*, *14*(2), 374–380. https://doi.org/10.1111/1755-0998.12180

Pinto, I., Nogueira, S., Rodrigues, S., Formigo, N., & Antunes, S. C. (2023). Can Zooplankton Add Value to Monitoring Water Quality? A Case Study of a Meso/Eutrophic Portuguese Reservoir. *Water*, *15*(9), 1678. https://doi.org/10.3390/w15091678

Põlme, S., Abarenkov, K., Henrik Nilsson, R., Lindahl, B. D., Clemmensen, K. E., Kauserud, H., Nguyen, N., Kjøller, R., Bates, S. T., Baldrian, P., Frøslev, T. G., Adojaan, K., Vizzini, A., Suija, A., Pfister, D., Baral, H.-O., Järv, H., Madrid, H., Nordén, J., … Tedersoo, L. (2020). FungalTraits: a user-friendly traits database of fungi and fungus-like stramenopiles. *Fungal Diversity*, *105*(1), 1–16. https://doi.org/10.1007/s13225-020-00466-2

Pond Action. (2002). *A guide to monitoring the ecological quality of ponds and canals using PSYM*. Pond Action (now Freshwater Habitats Trust).

Pond Conservation. (2009). *Pond Creation Toolkit - Freshwater Habitats Trust*. Million Ponds Project. https://freshwaterhabitats.org.uk/projects/million-ponds/pond-creation-toolkit/

Potvin, M., Rautio, M., & Lovejoy, C. (2022). Freshwater Microbial Eukaryotic Core Communities, Open-Water and Under-Ice Specialists in Southern Victoria Island Lakes (Ekaluktutiak, NU, Canada). *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.786094

Prasitwuttisak, W., Hoshiko, Y., Maeda, T., Haraguchi, A., & Yanagawa, K. (2022). Microbial Community Structures and Methanogenic Functions in Wetland Peat Soils. *Microbes and Environments*, *37*(3), ME22004. https://doi.org/10.1264/jsme2.ME22004

Prosser, J. I. (2020). Putting science back into microbial ecology: a question of approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/RSTB.2019.0240

Quaresma, A., Ankenbrand, M. J., Garcia, C. A. Y., Rufino, J., Honrado, M., Amaral, J., Brodschneider, R., Brusbardis, V., Gratzer, K., Hatjina, F., Kilpinen, O., Pietropaoli, M., Roessink, I., van der Steen, J., Vejsnæs, F., Pinto, M. A., & Keller, A. (2024). Semi-automated sequence curation for reliable reference datasets in ITS2 vascular plant DNA (meta-)barcoding. *Scientific Data*, *11*(1), 129. https://doi.org/10.1038/s41597-024-02962-5

R Core Team. (2018). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. .

RATNASINGHAM, S., & HEBERT, P. D. N. (2007). <scp>bold</scp> : The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Read, D. S., Gweon, H. S., Bowes, M. J., Newbold, L. K., Field, D., Bailey, M. J., & Griffiths, R. I. (2015). Catchment-scale biogeography of riverine bacterioplankton. *ISME Journal*, *9*(2), 516–526. https://doi.org/10.1038/ismej.2014.166

Rees, H., Baker, C. B., & Maddison, B. C. (2023). *An evidence review for great crested newt eDNA monitoring protocols Natural England Commissioned Report NECR476 www.gov.uk/natural-england*. www.gov.uk/natural-england

Rees, H. C., Baker, C. A., Gardner, D. S., Maddison, B. C., & Gough, K. C. (2017). The detection of great crested newts year round via environmental DNA analysis. *BMC Research Notes*, *10*(1), 327. https://doi.org/10.1186/s13104-017-2657-y

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, *94*(3), 849–873. https://doi.org/10.1111/brv.12480

Reji Chacko, M., Altermatt, F., Fopp, F., Guisan, A., Keggin, T., Lyet, A., Rey, P.-L., Richards, E., Valentini, A., Waldock, C., & Pellissier, L. (2023). Catchment-based sampling of river eDNA integrates terrestrial and aquatic biodiversity of alpine landscapes. *Oecologia*, *202*(4), 699–713. https://doi.org/10.1007/s00442-023-05428-4

Rengefors, K., Kremp, A., Reusch, T. B. H., & Wood, A. A. M. (2017). HORIZONS Genetic diversity and evolution in eukaryotic phytoplankton: revelations from population genetic studies. *J. Plankton Res*, *39*(2), 165–179. https://doi.org/10.1093/plankt/fbw098

Révéret, A., Rijal, D. P., & Alsos, I. G. (2023). *Environmental DNA of aquatic macrophytes: the potential for reconstructing past and present vegetation and environments*.

Reyne, M., Naaum, A. M., Marnell, F., Reid, N., & Helyar, S. J. (2021). Development and validation of a quantitative qPCR assay for detecting Natterjack toad (Epidalea calamita) eDNA samples. *Conservation Genetics Resources*, *13*(3), 319–322. https://doi.org/10.1007/S12686-021-01199-3/TABLES/1

Reyne, M., Nolan, M., McGuiggan, H., Aubry, A., Emmerson, M., Marnell, F., & Reid, N. (2021). Artificial agri-environment scheme ponds do not replicate natural environments despite higher aquatic and terrestrial invertebrate richness and abundance. *Journal of Applied Ecology*, *58*(2), 304–315. https://doi.org/10.1111/1365-2664.13738

Riaz, M., Kuemmerlen, M., Wittwer, C., Cocchiararo, B., Khaliq, I., Pfenninger, M., & Nowak, C. (2020). Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. *Ecological Applications*, *30*(2). https://doi.org/10.1002/eap.2034

Richardson, D. C., Holgerson, M. A., Farragher, M. J., Hoffman, K. K., King, K. B. S., Alfonso, M. B., Andersen, M. R., Cheruveil, K. S., Coleman, K. A., Farruggia, M. J., Fernandez, R. L., Hondula, K. L., López Moreira Mazacotte, G. A., Paul, K., Peierls, B. L., Rabaey, J. S., Sadro, S., Sánchez, M. L., Smyth, R. L., & Sweetman, J. N. (2022). A functional definition to distinguish ponds from lakes and wetlands. *Scientific Reports*, *12*(1), 10472. https://doi.org/10.1038/s41598-022-14569-0

Richardson, S. J., Clayton, R., Rance, B. D., Broadbent, H., McGlone, M. S., & Wilmshurst, J. M. (2015). Small wetlands are critical for safeguarding rare and threatened plant species. *Applied Vegetation Science*, *18*(2), 230–241. https://doi.org/10.1111/avsc.12144

Riley, W. D., Potter, E. C. E., Biggs, J., Collins, A. L., Jarvie, H. P., Jones, J. I., Kelly-Quinn, M., Ormerod, S. J., Sear, D. A., Wilby, R. L., Broadmeadow, S., Brown, C. D., Chanin, P., Copp, G. H., Cowx, I. G., Grogan, A., Hornby, D. D., Huggett, D., Kelly, M. G., … Siriwardena, G. M. (2018). Small Water Bodies in Great Britain and Ireland: Ecosystem function, human-generated degradation, and options for restorative action. In *Science of the Total Environment* (Vol. 645, pp. 1598–1616). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2018.07.243

Riva, F., & Fahrig, L. (2022). The disproportionately high value of small patches for biodiversity conservation. *Conservation Letters*, *15*(3), e12881. https://doi.org/10.1111/CONL.12881

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

Rivera, S. F., Vasselon, V., Jacquet, S., Bouchez, A., Ariztegui, D., & Rimet, F. (2018). Metabarcoding of lake benthic diatoms: from structure assemblages to ecological assessment. *Hydrobiologia*, *807*(1), 37–51. https://doi.org/10.1007/s10750-017-3381-2

Robert, M., & Hijmans, J. (2023). *Package "terra" Type Package Title Spatial Data Analysis*. https://github.com/rspatial/terra/issues/

Robinson, C. V., Uren Webster, T. M., Cable, J., James, J., & Consuegra, S. (2018). Simultaneous detection of invasive signal crayfish, endangered white-clawed crayfish and the crayfish plague pathogen using environmental DNA. *Biological Conservation*, *222*, 241–252. https://doi.org/10.1016/j.biocon.2018.04.009

Robinson, R. A., & Sutherland, W. J. (2002). Post-war changes in arable farming and biodiversity in Great Britain. *Journal of Applied Ecology*, *39*(1), 157–176. https://doi.org/10.1046/j.1365-2664.2002.00695.x

Robson, H. L. A., Noble, T. H., Saunders, R. J., Robson, S. K. A., Burrows, D. W., & Jerry, D. R. (2016). Fine-tuning for the tropics: application of <scp>eDNA</scp> technology for invasive fish detection in tropical freshwater ecosystems. *Molecular Ecology Resources*, *16*(4), 922–932. https://doi.org/10.1111/1755-0998.12505

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., De Wit, C. A., Hughes, T., Van Der

Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Liverman, D. (2009). Planetary Boundaries Exploring the Safe Operating Space for Humanity. In *Source: Ecology and Society* (Vol. 14, Issue 2). https://www.jstor.org/stable/26268316?seq=1&cid=pdf-reference#references_tab_contents

Rosentreter, J. A., Borges, A. V., Deemer, B. R., Holgerson, M. A., Liu, S., Song, C., Melack, J., Raymond, P. A., Duarte, C. M., Allen, G. H., Olefeldt, D., Poulter, B., Battin, T. I., & Eyre, B. D. (2021). Half of global methane emissions come from highly variable aquatic ecosystem sources. *Nature Geoscience*, *14*(4), 225–230. https://doi.org/10.1038/S41561-021-00715-2

Rosero, P., Crespo-Pérez, V., Espinosa, R., Andino, P., Barragán, Á., Moret, P., Gobbi, M., Ficetola, G. F., Jaramillo, R., Muriel, P., Anthelme, F., Jacobsen, D., Dangles, O., Condom, T., Gielly, L., Poulenard, J., Rabatel, A., Basantes, R., Cáceres Correa, B., & Cauvy-Fraunié, S. (2021). Multi-taxa colonisation along the foreland of a vanishing equatorial glacier. *Ecography*, *44*(7), 1010–1021. https://doi.org/10.1111/ecog.05478

Rourke, M. L., Fowler, A. M., Hughes, J. M., Broadhurst, M. K., DiBattista, J. D., Fielder, S., Wilkes Walburn, J., & Furlan, E. M. (2022). Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA*, *4*(1), 9–33. https://doi.org/10.1002/edn3.185

Royal Society for the Protection of Birds. (2008). *A 50-YEAR VISION FOR WETLANDS England's Wetland Landscape: securing a future for nature, people and the historic environment*.

Ruggeri, P., Pasternak, E., & Okamura, B. (2019). To remain or leave: Dispersal variation and its genetic consequences in benthic freshwater invertebrates. *Ecology and Evolution*, *9*(21), 12069–12088. https://doi.org/10.1002/ece3.5656

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. In *Global Ecology and Conservation* (Vol. 17, p. e00547). Elsevier B.V. https://doi.org/10.1016/j.gecco.2019.e00547

Ruuskanen, M. O., St. Pierre, K. A., St. Louis, V. L., Aris-Brosou, S., & Poulain, A. J. (2018). Physicochemical Drivers of Microbial Community Structure in Sediments of Lake Hazen, Nunavut, Canada. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.01138

Sadeghi, J., Chaganti, S. R., Shahraki, A. H., & Heath, D. D. (2021). Microbial community and abiotic effects on aquatic bacterial communities in north temperate lakes. *Science of The Total Environment*, *781*, 146771. https://doi.org/10.1016/j.scitotenv.2021.146771

Sagova-Mareckova, M., Boenigk, J., Bouchez, A., Cermakova, K., Chonova, T., Cordier, T., Eisendle, U., Elersek, T., Fazi, S., Fleituch, T., Frühe, L., Gajdosova, M., Graupner, N., Haegerbaeumer, A., Kelly, A. M., Kopecky, J., Leese, F., Nõges, P., Orlic, S., … Stoeck, T. (2021). Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Research*, *191*, 116767. https://doi.org/10.1016/J.WATRES.2020.116767

Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., Wangensteen, O. S., Baillie, C., Bryce, E., Dawson, D. A., Ochu, E., Hänfling, B., Handley, L. L., Mariani, S., Lambin, X., Sutherland, C., & McDevitt, A. D. (2019). Fishing for mammals: landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from lotic ecosystems. *BioRxiv*, 629758. https://doi.org/10.1101/629758

Sato, H., Sogo, Y., Doi, H., & Yamanaka, H. (2017). Usefulness and limitations of sample pooling for environmental DNA metabarcoding of freshwater fish communities. *Scientific Reports*, *7*(1), 14860. https://doi.org/10.1038/s41598-017-14978-6

Sawatzky, M. E., & Fahrig, L. (2019). Wetland buffers are no substitute for landscape-scale conservation. *Ecosphere*, *10*(4), e02661. https://doi.org/10.1002/ecs2.2661

Sawatzky, M. E., Martin, A. E., & Fahrig, L. (2019a). Landscape context is more important than wetland buffers for farmland amphibians. *Agriculture, Ecosystems and Environment*, *269*, 97–106. https://doi.org/10.1016/j.agee.2018.09.021

Sawatzky, M. E., Martin, A. E., & Fahrig, L. (2019b). Landscape context is more important than wetland buffers for farmland amphibians. *Agriculture, Ecosystems and Environment*, *269*, 97–106. https://doi.org/10.1016/j.agee.2018.09.021

Sayer, C., Andrews, K., Shilland, E., Edmonds, N., Edmonds-Brown, R., Patmore, I., Emson, D., & Axmacher, J. (2012). The role of pond management for biodiversity conservation in an agricultural landscape. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *22*(5), 626–638. https://doi.org/10.1002/aqc.2254

Sayer, C. D. (2014). Conservation of aquatic landscapes: ponds, lakes, and rivers as integrated systems. *Wiley Interdisciplinary Reviews: Water*, *1*(6), 573–585. https://doi.org/10.1002/wat2.1045

Sayer, C., Shilland, E., Greaves, H., & Dawson, B. , P. I. , E. D. , A. E. , R. P. , A. K. , A. J. and W. E. ,. (2013). Managing Britain's ponds-conservation lessons from a Norfolk farm. *British Wildlife*, *25*.

Scheffer, M., van Geest, G. J., Zimmer, K., Jeppesen, E., Søndergaard, M., Butler, M. G., Hanson, M. A., Declerck, S., & De Meester, L. (2006). Small habitat size and isolation can promote species richness: second-order effects on biodiversity in shallow lakes and ponds. *Oikos*, *112*(1), 227–231. https://doi.org/10.1111/j.0030-1299.2006.14145.x

Schenekar, T. (2023). The current state of eDNA research in freshwater ecosystems: are we shifting from the developmental phase to standard application in biomonitoring? *Hydrobiologia*, *850*(6), 1263–1282. https://doi.org/10.1007/s10750-022-04891-z

Schenekar, T., Schletterer, M., Lecaudey, L. A., & Weiss, S. J. (2020). Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: Lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. *River Research and Applications*, *36*(7), 1004–1013. https://doi.org/10.1002/rra.3610

Schloss, P. D. (2021). Amplicon Sequence Variants Artificially Split Bacterial Genomes into Separate Clusters. *MSphere*, *6*(4). https://doi.org/10.1128/mSphere.00191-21

Schloss, P. D. (2023). Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses. *BioRxiv*, 2023.06.23.546313. https://doi.org/10.1101/2023.06.23.546313

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*(9), e0162493. https://doi.org/10.1371/journal.pone.0162493

Scriver, M., Marinich, A., Wilson, C., & Freeland, J. (2015). Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. *Aquatic Botany*, *122*, 27–31. https://doi.org/10.1016/j.aquabot.2015.01.003

Seeber, P. A., Mcewen, G. K., Löber, U., Förster, D. W., Marion, |, East, L., Melzheimer, J., & Greenwood, A. D. (2019). Terrestrial mammal surveillance using hybridization capture of environmental DNA from African waterholes. | *Mol Ecol Resour*, *19*, 1486–1496. https://doi.org/10.1111/1755-0998.13069

Segarra, K. E. A., Schubotz, F., Samarkin, V., Yoshinaga, M. Y., Hinrichs, K.-U., & Joye, S. B. (2015). High rates of anaerobic methane oxidation in freshwater wetlands reduce potential atmospheric methane emissions. *Nature Communications*, *6*(1), 7477. https://doi.org/10.1038/ncomms8477

Semlitsch, R. D., Walls, S. C., Barichivich, W. J., & O'Donnell, K. M. (2017). Extinction Debt as a Driver of Amphibian Declines: An Example with Imperiled Flatwoods Salamanders. *Journal of Herpetology*, *51*(1), 12–18. https://doi.org/10.1670/16-090

Serrana, J. M., Miyake, Y., Gamboa, M., & Watanabe, K. (2019). Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity assessment and monitoring. *Ecological Indicators*, *101*, 963–972. https://doi.org/10.1016/j.ecolind.2019.02.008

Seviour, R. J., Mino, T., & Onuki, M. (2003). The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiology Reviews*, *27*(1), 99–127. https://doi.org/10.1016/S0168-6445(03)00021-4

Shackleton, M. E., Rees, G. N., Watson, G., Campbell, C., & Nielsen, D. (2019). Environmental DNA reveals landscape mosaic of wetland plant communities. *Global Ecology and Conservation*, *19*, e00689. https://doi.org/10.1016/j.gecco.2019.e00689

Shafi, S., Kamili, A. N., Shah, M. A., Parray, J. A., & Bandh, S. A. (2017). Aquatic bacterial diversity: Magnitude, dynamics, and controlling factors. In *Microbial Pathogenesis* (Vol. 104, pp. 39–47). Academic Press. https://doi.org/10.1016/j.micpath.2017.01.016

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation*, *197*, 131–138. https://doi.org/10.1016/j.biocon.2016.03.010

Shirazi, S., Meyer, R. S., & Shapiro, B. (2021). Revisiting the effect of PCR replication and sequencing depth on biodiversity metrics in environmental DNA metabarcoding. *Ecology and Evolution*, *11*(22), 15766–15779. https://doi.org/10.1002/ece3.8239

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., Pedersen, M. W., Jaidah, M. Al, Orlando, L., Willerslev, E., Møller, P. R., & Thomsen, P. F. (2016). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology & Evolution*, *1*(1), 0004. https://doi.org/10.1038/s41559-016-0004

Simberloff, D. S., & Wilson, E. O. (1970). Experimental Zoogeography of Islands. A Two-Year Record of Colonization. *Ecology*, *51*(5), 934–937. https://doi.org/10.2307/1933995

Simon, M., López-García, P., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P., & Jardillier, L. (2015). Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *The ISME Journal*, *9*(9), 1941–1953. https://doi.org/10.1038/ismej.2015.6

Smith, L. P., Clarke, L. E., Weldon, L., & Robson, H. J. (2022). An evidence-based study mapping the decline in freshwater ponds in the Severn Vale catchment in the UK between 1900 and 2019. *Hydrobiologia*, *849*(21), 4637–4649. https://doi.org/10.1007/s10750-022-05000-w

Smith, V. H., Foster, B. L., Grover, J. P., Holt, R. D., Leibold, M. A., & deNoyelles, F. (2005). Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. *Proceedings of the National Academy of Sciences*, *102*(12), 4393–4396. https://doi.org/10.1073/pnas.0500094102

Smucker, N. J., Pilgrim, E. M., Nietch, C. T., Darling, J. A., & Johnson, B. R. (2020). DNA metabarcoding effectively quantifies diatom responses to nutrients in streams. *Ecological Applications*, *30*(8). https://doi.org/10.1002/eap.2205

Soininen, J. (2007). ENVIRONMENTAL AND SPATIAL CONTROL OF FRESHWATER DIATOMS—A REVIEW. *Diatom Research*, *22*(2), 473–490. https://doi.org/10.1080/0269249X.2007.9705724

Soininen, J. (2012). Macroecology of unicellular organisms – patterns and processes. *Environmental Microbiology Reports*, *4*(1), 10–22. https://doi.org/10.1111/j.1758-2229.2011.00308.x

Soininen, J., Korhonen, J. J., Karhu, J., & Vetterli, A. (2011). Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. *Limnology and Oceanography*, *56*(2), 508–520. https://doi.org/10.4319/lo.2011.56.2.0508

Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J. J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J. C., Mooij, W. M., van Donk, E., & Winder, M. (2012). Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 429–448. https://doi.org/10.1146/annurev-ecolsys-110411-160251

SØNDERGAARD, M., JOHANSSON, L. S., LAURIDSEN, T. L., JØRGENSEN, T. B., LIBORIUSSEN, L., & JEPPESEN, E. (2010a). Submerged macrophytes as indicators of the ecological quality of

lakes. *Freshwater Biology*, *55*(4), 893–908. https://doi.org/10.1111/j.1365-2427.2009.02331.x

SØNDERGAARD, M., JOHANSSON, L. S., LAURIDSEN, T. L., JØRGENSEN, T. B., LIBORIUSSEN, L., & JEPPESEN, E. (2010b). Submerged macrophytes as indicators of the ecological quality of lakes. *Freshwater Biology*, *55*(4), 893–908. https://doi.org/10.1111/j.1365-2427.2009.02331.x

Soons, M. B., Brochet, A.-L., Kleyheeg, E., & Green, A. J. (2016). Seed dispersal by dabbling ducks: an overlooked dispersal pathway for a broad spectrum of plant species. *Journal of Ecology*, *104*(2), 443–455. https://doi.org/10.1111/1365-2745.12531

Spear, M. J., Embke, H. S., Krysan, P. J., & Vander Zanden, M. J. (2021). Application of eDNA as a tool for assessing fish population abundance. *Environmental DNA*, *3*(1), 83–91. https://doi.org/10.1002/edn3.94

Stefanidis, K., & Papastergiadou, E. (2019). Linkages between Macrophyte Functional Traits and Water Quality: Insights from a Study in Freshwater Lakes of Greece. *Water*, *11*(5), 1047. https://doi.org/10.3390/w11051047

Stein, E. D., Martinez, M. C., Stiles, S., Miller, P. E., & 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*(4), e95525. https://doi.org/10.1371/journal.pone.0095525

Stewart, G. A., & Robinson, C. (2017). Indoor and Outdoor Allergens and Pollutants. In *Middleton's Allergy Essentials* (pp. 73–116). Elsevier. https://doi.org/10.1016/B978-0-323-37579-5.00004-0

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

Stoneworts: their habitats, ecological requirements and conservation Integrated catchment science programme Science report: SC030202. (n.d.). Retrieved May 19, 2023, from www.environment-agency.gov.uk

Stoof-Leichsenring, K. R., Pestryakova, L. A., Epp, L. S., & Herzschuh, U. (2020). Phylogenetic diversity and environment form assembly rules for Arctic diatom genera—A study on recent and ancient sedimentary DNA. *Journal of Biogeography*. https://doi.org/10.1111/jbi.13786

Strategic Framework and Guidelines for the Future Development of the List of Wetlands of International Importance of the Convention on Wetlands (Ramsar, Iran, 1971) 2022 Update (2022).

Sun, J., Hunter, P. D., Tyler, A. N., & Willby, N. J. (2019). Lake and catchment-scale determinants of aquatic vegetation across almost 1,000 lakes and the contrasts between lake types. *Journal of Biogeography*, *46*(5), 1066–1082. https://doi.org/10.1111/jbi.13557

Symons, C. C., Pedruski, M. T., Arnott, S. E., & Sweetman, J. N. (2014). Spatial, Environmental, and Biotic Determinants of Zooplankton Community Composition in Subarctic Lakes and Ponds in Wapusk National Park, Canada. *Arctic, Antarctic, and Alpine Research*, *46*(1), 159–190. https://doi.org/10.1657/1938-4246-46.1.159

Szoszkiewicz, K., Jusik, S., Pietruczuk, K., & Gebler, D. (2019). The Macrophyte Index for Rivers (MIR) as an Advantageous Approach to Running Water Assessment in Local Geographical Conditions. *Water*, *12*(1), 108. https://doi.org/10.3390/w12010108

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For biodiversity research and monitoring. In *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Tai, X., Li, R., Zhang, B., Yu, H., Kong, X., Bai, Z., Deng, Y., Jia, L., & Jin, D. (2020). Pollution Gradients Altered the Bacterial Community Composition and Stochastic Process of Rural Polluted Ponds. *Microorganisms*, *8*(2), 311. https://doi.org/10.3390/microorganisms8020311

Talling, J. F. (1951). The element of chance in pond populations. *The Naturalist*, *839*, 157–170. https://www.biodiversitylibrary.org/item/201079#page/184/mode/1up

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., & Fontaneto, D. (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proceedings of the National Academy of Sciences*, *109*(40), 16208–16212. https://doi.org/10.1073/pnas.1209160109

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

Taylor, P. D., Fahrig, L., Henein, K., & Merriam, G. (1993). Connectivity Is a Vital Element of Landscape Structure. *Oikos*, *68*(3), 571. https://doi.org/10.2307/3544927

Taylor, P. D., Fahrig, L., & With, K. A. (2006). Landscape connectivity: a return to the basics. In *Connectivity Conservation* (pp. 29–43). Cambridge University Press. https://doi.org/10.1017/CBO9780511754821.003

Taylor, S., Gilbert, P. J., Cooke, D. A., Deary, M. E., & Jeffries, M. J. (2019). High carbon burial rates by small ponds in the landscape. *Frontiers in Ecology and the Environment*, *17*(1), 25–31. https://doi.org/10.1002/fee.1988

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L., & Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, *21*(11), 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. In *Biological Conservation* (Vol. 183, pp. 4–18). Elsevier Ltd. https://doi.org/10.1016/j.biocon.2014.11.019

Tian, Y., Lv, C., Huang, L., Shan, H., Wang, H., Wen, Z., Yin, C., Chou, Q., Zhang, X., Ni, L., & Cao, T. (2023). Seasonal variation and nutrient jointly drive the community structure of macrophytes in lakes with different trophic states. *Frontiers in Marine Science*, *10*. https://doi.org/10.3389/fmars.2023.1182823

Tilman, D., May, R. M., Lehman, C. L., & Nowak, M. A. (1994). Habitat destruction and the extinction debt. *Nature*, *371*(6492), 65–66. https://doi.org/10.1038/371065a0

Tornero, I., Boix, D., Bagella, S., Pinto-Cruz, C., Caria, M. C., Belo, A., Lumbreras, A., Sala, J., Compte, J., & Gascón, S. (2018). Dispersal mode and spatial extent influence distance-decay patterns in pond metacommunities. *PLOS ONE*, *13*(8), e0203119. https://doi.org/10.1371/journal.pone.0203119

Tournebize, J., Chaumont, C., & Mander, Ü. (2017). Implications for constructed wetlands to mitigate nitrate and pesticide pollution in agricultural drained watersheds. *Ecological Engineering*, *103*, 415–425. https://doi.org/10.1016/j.ecoleng.2016.02.014

Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, *183*, 93–102. https://doi.org/10.1016/j.biocon.2014.11.017

Tveit, A. T., Söllinger, A., Rainer, E. M., Didriksen, A., Hestnes, A. G., Motleleng, L., Hellinger, H.-J., Rattei, T., & Svenning, M. M. (2023). Thermal acclimation of methanotrophs from the genus Methylobacter. *The ISME Journal*, *17*(4), 502–513. https://doi.org/10.1038/s41396-023-01363-7

Urakawa, H., & Bernhard, A. E. (2017). Wetland management using microbial indicators. *Ecological Engineering*, *108*, 456–476. https://doi.org/10.1016/j.ecoleng.2017.07.022

US EPA. (2022). *Great Lakes Biology Monitoring Programme*.

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M., & Miya, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources*, *17*(6), e63–e75. https://doi.org/10.1111/1755-0998.12690

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., … Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular EcologyValentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., … Dejean, T. (2016). Next-Generation Monitoring of Aquatic Biodiversity Using Environmental DNA Metabarcoding. Molecular Ecology, 25(4), 929–942. Https://Doi.*, *25*(4), 929–942. https://doi.org/10.1111/mec.13428

Van de Meutter, F., Meester, L. De, & Stoks, R. (2007). METACOMMUNITY STRUCTURE OF POND MACROINVERTEBRATES: EFFECTS OF DISPERSAL MODE AND GENERATION TIME. *Ecology*, *88*(7), 1687–1695. https://doi.org/10.1890/06-0333.1

van der Gast, C. J. (2015). Microbial biogeography: The end of the ubiquitous dispersal hypothesis? In *Environmental Microbiology* (Vol. 17, Issue 3, pp. 544–546). Blackwell Publishing Ltd. https://doi.org/10.1111/1462-2920.12635

Van der Valk, A. (2012). *The Biology of Freshwater Wetlands*.

Vanschoenwinkel, B., Gielen, S., Vandewaerde, H., Seaman, M., & Brendonck, L. (2008). Relative importance of different dispersal vectors for small aquatic invertebrates in a rock pool metacommunity. *Ecography*, *31*(5), 567–577. https://doi.org/10.1111/j.0906-7590.2008.05442.x

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

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, *41*(18), 6396–6402. https://doi.org/10.1002/2014GL060641

Vimercati, G., Labadesse, M., Dejean, T., & Secondi, J. (2020). Assessing the effect of landscape features on pond colonisation by an elusive amphibian invader using environmental DNA. *Freshwater Biology*, *65*(3), 502–513. https://doi.org/10.1111/fwb.13446

Visco, J. A., Apothéloz-Perret-Gentil, L., Cordonier, A., Esling, P., Pillet, L., & Pawlowski, J. (2015). Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. *Environmental Science & Technology*, *49*(13), 7597–7605. https://doi.org/10.1021/es506158m

Waite, D. W., Chuvochina, M., Pelikan, C., Parks, D. H., Yilmaz, P., Wagner, M., Loy, A., Naganuma, T., Nakai, R., Whitman, W. B., Hahn, M. W., Kuever, J., & Hugenholtz, P. (2020). Proposal to reclassify the proteobacterial classes Deltaproteobacteria and Oligoflexia, and the phylum Thermodesulfobacteria into four phyla reflecting major functional capabilities. *International Journal of Systematic and Evolutionary Microbiology*, *70*(11), 5972–6016. https://doi.org/10.1099/ijsem.0.004213

Walker, D. (1970). Direction and rate in some British post-glacial hydroseres. In *Studies in the vegetational history of the British Isles*.

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. A., Jansson, J. K., Gregory Caporaso, J., Fuhrman, J. A., Apprill, A., Knight, R., & Walters, C. W. (2015). *Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys crossmark*. https://doi.org/10.1128/mSystems.00009-15

Walton, R. E., Sayer, C. D., Bennion, H., & Axmacher, J. C. (2020). Open-canopy ponds benefit diurnal pollinator communities in an agricultural landscape: implications for farmland pond management. *Insect Conservation and Diversity*. https://doi.org/10.1111/icad.12452

Wang, Y., Fan, Q., Wang, D., Zou, W.-Q., Tang, D.-X., Hongthong, P., & Yu, H. (2022). Species Diversity and Virulence Potential of the Beauveria bassiana Complex and Beauveria

scarabaeidicola Complex. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.841604

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., & Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, *5*(1), 27. https://doi.org/10.1186/s40168-017-0237-y

Wickham, H. (2016). *ggplot2*. Springer New York. https://doi.org/10.1007/978-0-387-98141-3

Wilden, B., Traunspurger, W., & Geisen, S. (2021). Inventory of the benthic eukaryotic diversity in the oldest European lake. *Ecology and Evolution*, *11*(16), 11207–11215. https://doi.org/10.1002/ece3.7907

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Bråthen, K. A., Yoccoz, N., … Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, *506*(7486), 47–51. https://doi.org/10.1038/nature12921

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A., & Cooper, A. (2003). Diverse plant and animal genetic records from holocene and pleistocene sediments. *Science*, *300*(5620), 791–795. https://doi.org/10.1126/science.1084114

Williams, P. (2017). *Pinkhill Meadow wetland creation scheme 2016: ecological survey of a 26 year old pond complex*.

Williams, P. (2018). *What's happening to the quality of our best ponds? A resurvey of National Pond Survey sites after 24 years*.

Williams, P., Biggs, J., Barr, C. J., & Cummins, C. P. (1998). *Lowland Pond Survey 1996*.

Williams, P., Biggs, J., Crowe, A., Murphy, J., Nicolet, P., Weatherby, A., & Dunbar, M. (2010). *Ponds Report from 2007*.

Williams, P., Biggs, J., Stoate, C., Szczur, J., Brown, C., & Bonney, S. (2019). Nature based measures reverse catchment biodiversity loss and increase freshwater resilience in an agricultural landscape. *BioRxiv*, 672915. https://doi.org/10.1101/672915

Williams, P., Biggs, J., Stoate, C., Szczur, J., Brown, C., & Bonney, S. (2020). Nature based measures increase freshwater biodiversity in agricultural catchments. *Biological Conservation*, *244*, 108515. https://doi.org/10.1016/j.biocon.2020.108515

Williams, P., Whitfield, M., & Biggs, J. (2007). How can we make new ponds biodiverse? A case study monitored over 7 years. In *Pond Conservation in Europe* (pp. 137–148). Springer Netherlands. https://doi.org/10.1007/978-90-481-9088-1_12

Williams, P., Whitfield, M., Biggs, J., Bray, S., Fox, G., Nicolet, P., & Sear, D. (2004). Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in Southern England. *Biological Conservation*, *115*(2), 329–341. https://doi.org/10.1016/S0006-3207(03)00153-8

Winder, M., & Cloern, J. E. (2010). The annual cycles of phytoplankton biomass. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3215–3226. https://doi.org/10.1098/rstb.2010.0125

With, K. A. (2019). *Essentials of Landscape Ecology*. Oxford University PressOxford. https://doi.org/10.1093/oso/9780198838388.001.0001

Wood, J. R., Burge, O. R., Bolstridge, N., Bonner, K., Clarkson, B., Cole, T. L., Davis, C., Fergus, A., King, P., McKeown, M. M., Morse, C., Richardson, S. J., Robertson, H., & Wilmshurst, J. M. (2021). Vertical distribution of prokaryotes communities and predicted metabolic pathways in New Zealand wetlands, and potential for environmental DNA indicators of wetland condition. *PLOS ONE*, *16*(1), e0243363. https://doi.org/10.1371/journal.pone.0243363

WRIGHT, J. F., BLACKBURN, J. H., GUNN, R. J. M., FURSE, M. T., ARMITAGE, P. D., WINDER, J. M., SYMES, K. L., & MOSS, D. (1996). Macroinvertebrate frequency data for the RIVPACS III sites in Great Britain and their use in conservation evaluation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *6*(3), 141–167. https://doi.org/https://doi.org/10.1002/(SICI)1099-0755(199609)6:3<141::AID-AQC189>3.0.CO;2-Z

Wu, H., Li, Y., Zhang, W., Wang, C., Wang, P., Niu, L., Du, J., & Gao, Y. (2019). Bacterial community composition and function shift with the aggravation of water quality in a heavily polluted river. *Journal of Environmental Management*, *237*, 433–441. https://doi.org/10.1016/j.jenvman.2019.02.101

Wurzbacher, C., Fuchs, A., Attermeyer, K., Frindte, K., Grossart, H. P., Hupfer, M., Casper, P., & Monaghan, M. T. (2017). Shifts among eukaryota, bacteria, and archaea define the vertical organization of a lake sediment. *Microbiome*, *5*(1). https://doi.org/10.1186/S40168-017-0255-9

Wurzbacher, C., Warthmann, N., Bourne, E. C., Attermeyer, K., Allgaier, M., Powell, J. R., Detering, H., Mbedi, S., Grossart, H. P., & Monaghan, M. T. (2016). High habitat-specificity in fungal communities in oligo-mesotrophic, temperate Lake Stechlin (North-East Germany). *MycoKeys*, *16*, 17–44. https://doi.org/10.3897/mycokeys.16.9646

WWF. (2018). *Living planet report 2018*. 41. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:LIVING+PLANET+REPORT+2004#0

Xie, Y., Wang, J., Yang, J., Giesy, J. P., Yu, H., & Zhang, X. (2017a). Environmental DNA metabarcoding reveals primary chemical contaminants in freshwater sediments from different land-use types. *Chemosphere*, *172*, 201–209. https://doi.org/10.1016/j.chemosphere.2016.12.117

Xie, Y., Wang, J., Yang, J., Giesy, J. P., Yu, H., & Zhang, X. (2017b). Environmental DNA metabarcoding reveals primary chemical contaminants in freshwater sediments from different land-use types. *Chemosphere*, *172*, 201–209. https://doi.org/10.1016/j.chemosphere.2016.12.117

Xu, H., Zhang, S., Ma, G., Zhang, Y., Li, Y., & Pei, H. (2020). 18S rRNA gene sequencing reveals significant influence of anthropogenic effects on microeukaryote diversity and composition along a river-to-estuary gradient ecosystem. *Science of The Total Environment*, *705*, 135910. https://doi.org/10.1016/J.SCITOTENV.2019.135910

Yang, H., Du, H., Qi, H., Yu, L., Hou, X., Zhang, H., Li, J., Wu, J., Wang, C., Zhou, Q., & Wei, Q. (2021). Effectiveness assessment of using riverine water eDNA to simultaneously monitor the riverine and riparian biodiversity information. *Scientific Reports*, *11*(1), 24241. https://doi.org/10.1038/s41598-021-03733-7

Yang, J., & Zhang, X. (2020). eDNA metabarcoding in zooplankton improves the ecological status assessment of aquatic ecosystems. *Environment International*, *134*, 105230. https://doi.org/10.1016/j.envint.2019.105230

Yang, N., Wang, L., Lin, L., Li, Y., Zhang, W., Niu, L., Zhang, H., & Wang, L. (2022). Pelagic-benthic coupling of the microbial food web modifies nutrient cycles along a cascade-dammed river. *Frontiers of Environmental Science & Engineering*, *16*(4), 50. https://doi.org/10.1007/s11783-021-1484-5

Yang, Y., Gao, Y., Chen, Y., Li, S., & Zhan, A. (2019). Interactome-based abiotic and biotic impacts on biodiversity of plankton communities in disturbed wetlands. *Diversity and Distributions*, *25*(9), 1416–1428. https://doi.org/10.1111/ddi.12949

Yates, M. C., Glaser, D. M., Post, J. R., Cristescu, M. E., Fraser, D. J., & Derry, A. M. (2021). The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. *Molecular Ecology*, *30*(13), 3068–3082. https://doi.org/10.1111/mec.15543

Yi, Z., Berney, C., Hartikainen, H., Mahamdallie, S., Gardner, M., Boenigk, J., Cavalier-Smith, T., & Bass, D. (2017). High-throughput sequencing of microbial eukaryotes in Lake Baikal reveals ecologically differentiated communities and novel evolutionary radiations. *FEMS Microbiology Ecology*, *93*(8). https://doi.org/10.1093/femsec/fix073

Yuan, B., Guo, M., Wu, W., Zhou, X., Li, M., & Xie, S. (2023). Spatial and Seasonal Patterns of Sediment Bacterial Communities in Large River Cascade Reservoirs: Drivers, Assembly Processes, and Co-occurrence Relationship. *Microbial Ecology*, *85*(2), 586–603. https://doi.org/10.1007/S00248-022-01999-6/FIGURES/2

Zhang, J., Jiao, S., & Lu, Y. (2018a). Biogeographic distribution of bacterial, archaeal and methanogenic communities and their associations with methanogenic capacity in Chinese wetlands. *Science of The Total Environment*, *622–623*, 664–675. https://doi.org/10.1016/j.scitotenv.2017.11.279

Zhang, J., Jiao, S., & Lu, Y. (2018b). Biogeographic distribution of bacterial, archaeal and methanogenic communities and their associations with methanogenic capacity in Chinese wetlands. *Science of The Total Environment*, *622–623*, 664–675. https://doi.org/10.1016/j.scitotenv.2017.11.279

Zhang, L., Xu, S., Zhao, Z., Zhou, X., & Feng, Q. (2021). Precision of eDNA Metabarcoding Technology for Biodiversity Monitoring of Eukaryotic Phytoplankton in Lakes. *Huan Jing Ke Xue.* , *8*(42(2)), 796–807.

Zhang, L., Zhao, T., Shen, T., & Gao, G. (2019). Seasonal and spatial variation in the sediment bacterial community and diversity of Lake Bosten, China. *Journal of Basic Microbiology*, *59*(2), 224–233. https://doi.org/10.1002/jobm.201800452

Zhang, N., Xiao, X., Pei, M., Liu, X., & Liang, Y. (2017). Discordant Temporal Turnovers of Sediment Bacterial and Eukaryotic Communities in Response to Dredging: Nonresilience and Functional Changes. *Applied and Environmental Microbiology*, *83*(1). https://doi.org/10.1128/AEM.02526-16

Zhang, Y., Pavlovska, M., Stoica, E., Prekrasna, I., Yang, J., Slobodnik, J., Zhang, X., & Dykyi, E. (2020). Holistic pelagic biodiversity monitoring of the Black Sea via eDNA metabarcoding approach: From bacteria to marine mammals. *Environment International*, *135*, 105307. https://doi.org/10.1016/j.envint.2019.105307

Zheng, B.-H., Chen, Z.-J., Li, Y.-Y., Fohrer, N., Zhang, Y., Wu, D.-Y., Yan, X.-Y., & Li, B.-L. (2020). Structural Characteristics and Driving Factors of the Planktonic Eukaryotic Community in the Danjiangkou Reservoir, China. *Water*, *12*(12), 3499. https://doi.org/10.3390/w12123499

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*(3), 526–542. https://doi.org/10.1111/1755-0998.12336

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L., Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. https://doi.org/10.1111/mec.15060