UNIVERSITY OF HULL

Environmental DNA (eDNA) monitoring of priority conservation fish species in UK lentic ecosystems

being a Thesis submitted for the Degree of Ph.D.
in the University of Hull

by

Cristina Di Muri

BSc, MSc (Hons), University of Ancona

October 2020

To my dad,

dancing in the mind field

with you

Declaration of Authorship

I declare that the work herein is intellectually my own. Each data chapter received contribution from my supervisors (Bernd Hänfling, B.H.; Lori Lawson Handley, L.L.H.), advisors (Colin W. Bean, C.W.B.; Ian J. Winfield, I.J.W.), and collaborators (Colin E. Adams, C.E.A.; Ben Aston, B.A.; Marco Benucci, M.B.; Rein Brys, R.B.; Andy Gowans; A.G.; Nathan Griffith, N.G.; Lynsey R. Harper, L.R.H.; Ben J. James, B.J.J.; Jianlong Li, J.L.; Graeme Peirson, G.P.; Glenn Rhodes, G.R.; Philip A. Rippon, P.A.R.; Graham S. Sellers, G.S.S.; Peter Shum, P.S.; Kerry Walsh, K.W.; Hayley V. Watson; H.V.W.).

Chapter 2

I led the experiment which was designed by B.H., K.W. and G.P. H.V.W. assisted with sampling and initial samples processing. G.S.S. helped with optimization of DNA extraction. I performed all subsequent laboratory works and bioinformatics supervised by J.L. I analysed the data and wrote the first draft of the paper which was then revised under the advice of supervisors (B.H., L.L.H.), advisors (C.W.B., I.J.W.) and collaborators (K.W., G.P., H.V.W., J.L., G.S.S.).

Chapter 3

I led the experiment which was co-designed with B.H., I.J.W. and B.J.J. H.V.W, L.R.B., J.L., M.B., I.J.W., B.J.J. assisted with fieldwork and initial samples processing. I performed all subsequent laboratory works and bioinformatics. I analysed the data and wrote the first draft of the paper which was then revised under the advice of supervisors (B.H., L.L.H.).

Chapter 4

I led the experiment which was co-designed with B.H. M.B., J.L., I.J.W., B.J.J. assisted with fieldwork and initial samples processing. L.R.H. helped with the TaqMan qPCR assay design. G.R. and N.G. provided advice on qPCR experiments. R.B. provided the SYBRGreen qPCR protocol. I performed all laboratory works and bioinformatics. I analysed the data and wrote the first draft of the paper which was then revised under the advice of supervisors (B.H., L.L.H.).

Chapter 5

I used data collected from the University of Hull in collaboration with the Environment Agency, Natural Resources Wales and Scottish Environment Protection Agency. B.A., P.A.R., and A.G. contributed to some data collection. I assisted with sampling and initial sample processing. I analysed the data based on information provided by C.W.B, C.E.A. and I.J.W. and wrote the first draft of the paper which was then revised under the advice of supervisors (B.H., L.L.H.).

Acknowledgements

To my lab mates, Marco Benucci, Rosetta Blackman, Nathan Griffiths, Lynsey R. Harper, Jianlong Li, Graham S. Sellers, Peter Shum, Michael Winter and Hayley V. Watson. I am so grateful for your personal, scientific and psychological support that I cannot even quantify how much of this PhD I owe you (Marco I know you probably kept count of the £5!). I would not have made it through without you guys!

I would like to thank the EvoHull group Rob Donnelly, Africa Gómez, Domino Joyce, Dave Lunt and Katharina Wollenberg-Valero, whose invaluable source of knowledge and precious advice made me a better scientist and improved my research.

A special thanks goes to Stephanie McLean without which support, inside and outside the University, I would not have survived the last two years. You are among the best things this PhD has granted me!

I am so grateful for all people of the "Hull network and stuff" group that make the whole PhD experience in Hull just extraordinary. Listing all of you would be impossible, but without the little daily happy things that each of you has provided over the years, I would have struggled to keep my positivity at the top. A special thanks to Ana Maria for our daily walks to the University (and much more of course!), to Barbara who tried to educate me to a correct work-life balance (sorry if I miserably failed!), to Ola and our glasses of wine (waiting for the next cheer!), to Angelo for being my Italian family in Hull and finally to my beloved "husband" Pili, I am not sure where this unpredictable life will take us, but I know that we'll always be there for each other! Love you so much!

Time to move on to my Italian supporters, therefore, I shall switch to Italian.

Mamma grazie per il tuo continuo supporto che non è venuto a mancare mai, neanche negli ultimi anni che sono stati così difficili per noi. Grazie per avermi coccolata (con cibo e non solo!) e protetta sempre. Grazie alla mia famiglia ed in particolare alle zie (Teresa, Franca ed Ester) e a mia cugina Chiara per avermi aiutata ad affrontare questi ultimi duri mesi in assenza di una figura essenziale della mia vita.

Tanto Matti, la mia famiglia adottiva, rimarrete sempre una delle cose più preziose che ho. Il reciproco amore, la stima e il supporto che da anni ci uniscono sono sicura ci accompagneranno sempre in futuro e, nonostante le distanze, riusciremo a gioire reciprocamente per ogni prossimo traguardo.

Ultimo, solo in ordine di arrivo, grazie a Matteo che negli ultimi mesi ha portato qualcosa di inaspettato e bellissimo nella mia vita. Non avrei mai pensato di poter incontrare una persona così speciale da decidere di volerla accanto a me come compagno di questo viaggio che è la vita, ma ora ci sei, e non ti lascerò andare via tanto facilmente. Grazie per tutte le piacevoli distrazioni che mi hai regalato in questi mesi e per l'affetto immenso che giornalmente mi dimostri.

Papà questa tesi è dedicata a te e a tutti i nostri sogni, spero che tu possa essere orgoglioso di me come lo sei sempre stato. Grazie a te per avermi reso la persona che sono oggi, essere la tua "copia", come molti dicono, è ciò che mi rende più orgogliosa.

Table of contents

Abstract

Freshwater environments cover only a tiny fraction of the planet's surface, yet they are biodiversity hotspots and support almost half of the global fish diversity, including globally and locally endangered species. In the UK, three species of national conservation value namely, Arctic charr (*Salvelinus alpinus* L.), European whitefish (*Coregonus lavaretus* L.) and vendace (*Coregonus albula* L.), are threatened throughout their limited distribution range by a number of human-driven environmental changes.

To combat a further loss of these priority conservation fish, non-invasive, sensitive and reliable monitoring tools are required to assess their status and guide appropriate conservation measures. Advances in the field of environmental DNA (eDNA) have shown the potential of this tool to detect low-abundant species, however the actual application of eDNA methods (e.g. targeted and metabarcoding) for rare species monitoring has been mostly limited to the collection of presence/absence information.

The scope of this thesis is to explore novel aspects of eDNA metabarcoding including the ability to estimate fish biomass/abundance, determine fish spawning and assess distribution and abundance of priority conservation fish across a variety of UK lentic systems. This thesis provides evidence that eDNA metabarcoding accurately reflects absolute fish biomass and abundance in high fish density ponds, and quantitative estimates are repeatable between different methods of DNA capture. This thesis demonstrates that eDNA metabarcoding determines location, timing and effort of Arctic charr spawning and can be as sensitive as quantitative PCR. Lastly, eDNA metabarcoding provides accurate information of Arctic charr, vendace and European whitefish distribution and abundance in UK lentic systems including sites of special conservation interest.

The fish community information obtained from eDNA metabarcoding is, therefore, beneficial to predict future changes in distribution and abundance of priority conservation fish as well as to design appropriate management actions, and can enormously contribute to more efficient monitoring programmes in UK lentic systems.

Chapter 1

General introduction

1.1 Freshwater ecosystems

Freshwaters provide vital resources for both humans and ecosystem services, supporting around 9.5% of all known species while covering only 0.8% of the globe's surface (Dudgeon *et al.* 2006; Strayer and Dudgeon, 2010). With the rapid and constant increase of the human population and global economy, the worldwide overexploitation of freshwaters and their surroundings has a profound negative ecological impact on these fragile ecosystems, with these biodiversity hotspots being among the most jeopardized on Earth with an estimated decline far greater than their terrestrial counterparts (Ricciardi and Rasmussen 1999; Sala *et al.*, 2000). On the other hand, the fragmented nature of lentic freshwater habitats has led to the evolution of many endemisms, whereby fragmentation reduces i) the efficiency of species to disperse and re-establish local populations that have been extirpated and ii) their ability to respond to climate change, resulting in species or populations that are extremely sensitive to human stressors (see Collen *et al.*, 2013 for an overview). Multiple and interacting stresses such as overexploitation, organic and inorganic water pollution, destruction or degradation of habitats and non-native species introduction represent the major threats to freshwater diversity, that are primarily concentrated in the areas of intense agriculture, industry or domestic activity (Dudgeon *et al.*, 2000; Collen *et al.*, 2013). Human-induced climate change, acting through rising water temperatures and hydrological changes, further reduces the integrity and function of freshwater systems (Dudgeon *et al.*, 2006; Strayer and Dudgeon, 2010). In the face of these threats, the protection and sustainable management of freshwater ecosystems is critical (IUCN, 2019). To achieve informed management actions and plan efficient conservation efforts, effective monitoring programmes providing estimates of the system's state are crucial for structured decision-making (Nichols and Williams, 2006).

In freshwater environments, vertebrates represent a third of the total diversity described, including around 13,000 described fish species that account for 40%–45% of global fish diversity (Cooke *et al.*, 2012). Freshwater fishes provide ecosystem services

of major economic, nutritional, scientific, historical and cultural importance (IUCN FFSG, 2015), and are commonly used for evaluating the functioning and status of freshwater ecosystems and habitat quality highlighting changes in species diversity (e.g. European Water Framework Directive).

1.2 Freshwaters fish monitoring - from conventional methods to molecular investigation

A broad range of different methods have been used for the assessment of freshwater fish populations worldwide, but traditionally these methods rely on the isolation and morphological identification of the species monitored (Murphy *et al.*, 1996). Capture methods involve the physical removal of fish from water through netting/trapping using fyke nets, gill-nets and trawling, and more recently, electrofishing and hydroacoustic methods have been used to complement traditional capture methods (Winfield *et al.*, 2009). Capture and non-capture methods are usually combined for freshwater fish monitoring since no single method has universal applicability and has associated caveats (Winfield *et al.*, 2009). Capture methods may lack sufficient power to provide robust estimates of abundance, whereas non-capture methods may suffer from taxonomic ambiguity (Radinger *et al.*, 2019). Furthermore, logistical or ethical questions limit the employment of these methods. For example, gill-netting is associated with varying levels of acceptance in different countries resulting from its destructive features, while hydroacoustic methods are only manageable in easy-to-access locations (Winfield, 2002). Traditional biomonitoring is also costly, labour intensive, and requires taxonomic expertise for accurate species identification (Radinger *et al.*, 2019). Thus, the limitations of conventional fish monitoring have created a demand for alternative strategies and genetic methods, at the forefront of such strategies, are fast becoming the biggest "game-changer" in biodiversity monitoring (Lawson Handley, 2015).

Obtaining information about species and communities by retrieving DNA from environmental samples has the ability to overcome some of the challenges associated with established monitoring methods. The molecular investigation of environmental samples is known as environmental DNA (eDNA). This genetic tool has long been used

in microbiology (Ogram *et al.*, 1987), but its potential for application in other fields has only been recognised in the last 15 years. In 2005, Martellini *et al.* retrieved macroorganisms' DNA (i.e. human, pig, sheep and cow) from freshwater samples and, in 2008, Ficetola *et al.* demonstrated the potential application of this tool for biodiversity monitoring with the identification of an invasive amphibian in water samples.

Environmental DNA is a heterogeneous mixture of genetic material ranging from cells to freely floating fragments that can be isolated from an environmental sample such as water, soil, air or faeces (Taberlet *et al.*, 2012). Environmental DNA is ubiquitous as organisms shed their genetic material in the surroundings through metabolic waste, damaged tissues, sloughed skin cells and decomposition (Thomsen and Willerslev, 2015). The analysis of eDNA from an environmental sample consists of extraction of genetic material followed by a Polymerase Chain Reaction (PCR) to amplify target DNA. To identify a single species within a mixed environmental sample, the most common approach is to use species-specific markers and conventional PCR, quantitative PCR (qPCR) or digital droplet PCR (ddPCR) to detect unique DNA sequences of the target species. Alternatively, the use of next-generation sequencing (NGS) allows the simultaneous identification of many species within a taxonomic group. This community-wide approach is known as DNA metabarcoding and involves the use of broad-range primers during PCR that amplify a range of species. In recent years, the cost of this technology has reduced making this approach very attractive in conservation management and scientific research.

Despite the enthusiasm of the scientific community, it is believed that eDNA-based methods cannot replace established methods as molecular approaches also hold some limitations and are associated with some challenging technical aspects. For example, eDNA is unable to determine the age class or life stage of organisms, and cannot distinguish between living and deceased individuals (Lawson Handley, 2015). Additionally, eDNA is often degraded and occurs in low concentrations in the environment, thus eDNA is very prone to contamination, which leads to false positive detections (Type I error). False positives can also arise from amplification errors during PCR (low specificity of primers/probes; competition with non-target templates) or DNA

sources other than living animals present in the environment (wastewater, sewage, faeces from predatory animals). Furthermore, the presence of PCR inhibitors in the environment (i.e. humic substances) or low-concentrations of target DNA may lead to false negatives (Type II error). Imperfect detections can be largely addressed through optimised sampling effort and laboratory practices (Goldberg *et al.*, 2016), however there is no standard protocol suitable for all type of eDNA surveys, thus methods are usually chosen and adapted to the question, target species, type of environment and the experimental constraints.

1.3 Environmental DNA research in lentic systems

Freshwater ecosystems are among the most researched environments for eDNA studies on macroorganisms, most likely because biomonitoring in freshwater ecosystems is imposed by law in a growing number of countries (e.g. European Council, 2000) and, additionally, because eDNA methods have often outperformed established methods in biodiversity monitoring of freshwater systems (Jerde *et al.*, 2011; Pilliod *et al.*, 2013; Civade *et al.*, 2016; Hänfling *et al.*, 2016; Olds *et al.*, 2016; Valentini *et al.*, 2016).

The distribution and concentration of eDNA in freshwater ecosystems varies as a function of factors that influence the production, degradation, transport and persistence of DNA in water. The input rate of DNA in water differs across species and is influenced by density and biomass of organisms as well as their seasonal activity (Barnes and Turner, 2016; de Souza *et al.*, 2016). Conversely, several environmental factors such as high temperatures, UV-radiation and pH promote eDNA degradation acting through a variety of mechanisms (Barnes *et al.*, 2014; Strickler *et al.*, 2015). Biotic factors such as microorganisms and extracellular enzymes further contribute to eDNA degradation, and physical processes determine the transport, sedimentation and resuspension of eDNA in water (Turner *et al.*, 2015). Environmental DNA in aquatic systems decays exponentially after shedding (Dejean *et al.*, 2011; Thomsen *et al.*, 2012a; Piaggio *et al.*, 2014; Pilliod *et al.*, 2014), and this rapid degradation makes this tool useful for describing biodiversity since positive detections are likely to be associated with contemporary presence of species. The distribution of eDNA in lentic

systems additionally reflects the species' spatiotemporal patterns as a consequence of their habitat preferences, avoidance of inter and intra-specific competition and predation (Hänfling *et al.*, 2016; Sato *et al.*, 2017; Lawson Handley *et al.*, 2019; Li *et al.*, 2019a; Zhang *et al.*, 2019; Harper *et al.*, 2020). For example, Hänfling *et al.* (2016) observed an association between fish species eDNA and trophic conditions in England's largest lake, and Takahara *et al.*, (2012) found a patchy distribution of common carp eDNA in a lagoon whereby the concentration of eDNA was 40% lower in low-use areas as opposed to high-use areas, even when these areas were only tens or hundreds of meters apart.

In lentic systems, the spatial patterns of fish distribution and eDNA are known to vary across seasons. Temporal dynamics of eDNA fluctuate according to the species ecology as well as variation in environmental conditions influencing water mixing and different rates of DNA degradation throughout the year. In temperate lakes, during the summer months, the vertical water stratification determines a stronger pattern in eDNA spatial distribution compared to winter (Lawson Handley *et al.*, 2019). Thus, fish eDNA distribution in temperate lentic waters is heterogeneous over short distances both horizontally and vertically with more evident patterns in summer when water is not mixed (Takahara *et al.*, 2012; Lawson Handley *et al.*, 2019). In the warmer months, the spatial distribution of eDNA is also influenced by the higher degradation rate as a consequence of higher temperatures and ultraviolet radiations (Tsuji *et al.*, 2017; Li *et al.*, 2019b).

Despite the complex and multidimensional eDNA dynamics, eDNA-based approaches in lentic systems have been demonstrated to be an effective tool for monitoring low-density fish species such as invasive (Jerde *et al.*, 2011; Mahon *et al.*, 2013; Turner *et al.*, 2014) or endangered taxa including locally-threatened species (Janosik and Johnston, 2015; Eva *et al.*, 2016; Bylemans *et al.*, 2017; Mizumoto *et al.*, 2018), commercially important endangered fish (Laramie *et al.*, 2015; Griffiths *et al.*, 2020), and near extinct taxa (Sigsgaard *et al.*, 2015; Pfleger *et al.*, 2016).

1.4 Thesis rationale

This thesis aims to evaluate aspects of eDNA metabarcoding (e.g. accuracy and reliability for quantification, sensitivity, manageability on a broad-scale etc.) for the implementation of this monitoring tool to the assessment and management of priority conservation fish species in UK lentic systems. This thesis will focus on three specific species: Arctic charr (*Salvelinus alpinus* L. [chapter 3 to chapter 5]), vendace (*Coregonus albula* L. [chapter 5]) and European whitefish (*C. lavaretus* L. [chapter 5]). These three salmonids are listed in the UK Biodiversity Action Plan (BAP; see Table 1; https://jncc.gov.uk/our-work/uk-bap/) and are protected in several sites across their limited distribution range (Maitland *et al.*, 2007; Winfield *et al.*, 2012; Winfield *et al.*, 2013).

Table 1. Priority conservation fish species in the UK. The information was produced between 1995 and 1999 and updated in 2007 in response to the Species and Habitats Review Report (BRIG, 2007). Table reports scientific and common names as well as fish distribution across the UK where the symbol ✓ stands for presence and the symbol ✗ for absence. Modified from "List of UK BAP Priority Fish (2007)".

| SCIENTIFIC NAME	COMMON NAME	ENGLAND	SCOTLAND	WALES	NORTHERN IRELAND
Acipenser sturio	Common sturgeon	✓	✓	✗	✓
Alosa alosa	Allis shad	✓	✓	✓	✓
Alosa fallax	Twaite shad	✓	✓	✓	✓
Anguilla anguilla	European eel	✓	✓	✓	✓
Cobitis taenia	Spined loach	✓	✗	✗	✗
Coregonus albula	**Vendace**	✓	✓	✗	✗
Coregonus autumnalis	Pollan	✗	✗	✗	✓
Coregonus lavaretus	**European whitefish**	✓	✓	✓	✗
Lampetra fluviatilis	European river lamprey	✓	✓	✓	✓
Lota lota	Burbot	✗	✗	✗	✗
Osmerus eperlanus	Smelt (sparling)	✓	✓	✓	✓
Petromyzon marinus	Sea lamprey	✓	✓	✓	✓
Salmo salar	Atlantic salmon	✓	✓	✓	✓
Salmo trutta	Brown/sea trout	✓	✓	✓	✓
Salvelinus alpinus	**Arctic charr**	✓	✓	✓	✗
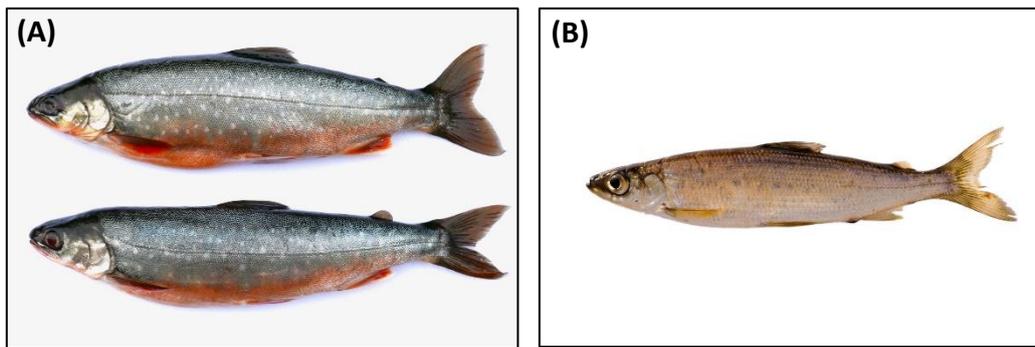
Figure 1. Arctic charr males (**A**) and vendace (**B**). Photo credit: Dr. Winfield I.J., used with permission.

Ecology and national conservation value of Arctic charr and coregonids in the UK

Arctic charr and coregonids (vendace and European whitefish) are key species in UK lentic systems and their conservation value spans from scientific to commercial and cultural interest (Winfield *et al.*, 2004; Winfield *et al.*, 2013; Winfield *et al.*, 2019). Anthropogenic pressures such as pollution, eutrophication, acidification, afforestation, engineering, non-native species introduction and climate change are jeopardising the species' native habitats, and many populations became extinct in several original sites (Maitland *et al.*, 2007; Maitland and Lyle, 2013). Vendace is the UK's rarest freshwater fish and it naturally exists in only two sites in the English Lake District, Derwent Water and Bassenthwaite Lake, and only the population from Derwent Water is considered stable (Winfield *et al.*, 2012). European whitefish distribution is also limited to seven native sites scattered between Scotland, England and Wales (Winfield *et al.*, 2013). The Arctic charr is the most broadly distributed species of these three salmonids, yet several populations are now extinct and many others may suffer the same fate in the near future (Maitland *et al.*, 2007). Arctic charr and coregonids dwell oligotrophic and deep lakes with low water temperatures, high concentration of dissolved oxygen and spawning areas with hard substrates and no overlying fine sediment (Winfield *et al.*, 2008a; Winfield *et al.*, 2012). These requirements limit the distribution range of these species and are the reasons for the species' extirpation in many sites as a consequence of habitat degradation. For example, nutrient enrichment leading to low oxygen levels in water (Winfield *et al.*, 2008a) or sedimentation reducing the number of suitable spawning areas in lakes (Miller *et al.*, 2015) or competition with introduced non-native species (Winfield *et al.*, 2008a). The widespread decline of these species in UK

freshwater habitats has required a number of conservation efforts consisting in habitat restoration or fish stock translocations in refuge sites (Maitland *et al.*, 2007; Maitland and Lyle, 2013). Despite the species protection and active management by national legislation (Bean 2003a; Bean 2003b; Winfield *et al.*, 2013) there is still a major concern towards many populations especially those which are poorly monitored given a lack of resources required for accurate, periodic and broad-scale species monitoring.

In this thesis we evaluate different aspects of eDNA metabarcoding to understand how the implementation of this non-invasive molecular method could contribute to the effective assessment of these threatened fish populations in UK lentic systems.

Can eDNA metabarcoding provide reliable estimates of fish biomass and abundance?

The collection of biomass and abundance data is one of the most commonly used metrics for fisheries management and conservation of low-abundant species. Quantitative data are difficult to obtain in aquatic ecosystems and so far established methods such as electrofishing and hydroacoustic have been used to provide at least semi-quantitative estimates of fish populations (e.g. Winfield *et al.*, 2009; Winfield *et al.*, 2015). Understanding the association between the amount of eDNA in water and species abundance would greatly enhance the power and the applications of this non-invasive monitoring tool. The existence of a relationship between the concentration of eDNA released from fish in water and fish biomass/abundance has been experimentally verified by qPCR in aquariums, mesocosms, and experimental ponds (Takahara *et al.*, 2012; Evans *et al.*, 2016), in lakes (Lacoursière-Roussel et al, 2016a; Klobucar *et al.*, 2017; Capo *et al.*, 2019), in streams and rivers (Pilliod *et al.*, 2013; Jane *et al.*, 2015; Baldigo *et al.*, 2017; Doi *et al.*, 2017; Itakura *et al.*, 2019), and in marine environments (Salter *et al.*, 2019; Takahashi *et al.*, 2020).

Although the correlation between fish biomass and fish eDNA concentrations based on qPCR seems straightforward, the situation becomes increasingly more complex when targeting a set of species using eDNA metabarcoding as several factors can influence quantitative estimates of community-wide eDNA approaches. For example, the amount of DNA released in water for a given biomass may vary among different species (Sassoubre *et al.*, 2016), the number of reads assigned to each species is

9

associated with the quality of the match between primers and their binding regions (Piñol *et al.*, 2019), and the competitive amplification may affect the original relative abundance of different DNA templates in a mixed sample (Kebschull and Zador, 2015). Despite these and many other complications (see also Chapter 2), a few studies have demonstrated the potential of eDNA metabarcoding to provide semi-quantitative estimates of fish biomass in artificial (Kelly *et al*, 2014; Evans *et al.*, 2016; Li *et al.*, 2018) and natural environments (Thomsen *et al.*, 2012b; Hänfling *et al.*, 2016; Sard *et al.*, 2019).

So far, in natural aquatic systems, estimates of fish biomass and abundance from eDNA metabarcoding have been compared to semi-quantitative and quantitative data obtained with conventional methods (e.g. catches, visual count and electrofishing). Despite the observed good associations between conventional and molecular methods, a certain scepticism among researchers and environmental managers is still present since both methods have their own biases and limitations. Chapter 2 explores the power of eDNA metabarcoding to provide reliable quantitative estimates of fish populations in semi-natural ponds with known fish biomass and abundance. Proving the existence of strong associations between eDNA metabarcoding data and actual measures of fish biomass and abundance in semi-natural freshwater environments represent a key step towards the integration of using eDNA in future monitoring plans.

Can eDNA metabarcoding determine the location and timing of fish spawning?

Safeguarding the species' reproductive activity is critical to maintain self-sustaining populations and, obtaining accurate information on the location and timing of fish spawning activity can benefit the effective conservation and management of fish populations. Assessment of gonad maturation, otolith micro-chemistry, radio tagging, scale and blood sample collection can be used to monitor fish reproductive activity (Bylemans *et al.*, 2018). However, these approaches can be harmful or result in the loss of individuals, so visual observations and acoustic surveys are often preferred for monitoring spawning of low-abundant fish populations (Edgar *et al.*, 2004; Miller *et al.*, 2012).

The use of eDNA to survey fish spawning can be a valuable and non-invasive alternative to conventional approaches. The act of spawning has been initially used to enhance detection probability from eDNA of low-abundance fish species (Janosik and Johnston, 2015; Eiler *et al.*, 2018), since the release of gametes in water is species-specific and results in a localised increase of eDNA. Thereafter it has been suggested that the absolute abundance of mitochondrial DNA in water can be used to determine the reproductive status of populations (Spear *et al.*, 2015; Bylemans *et al.*, 2018). A number of studies have used qPCR to assess localised and time-limited variation of eDNA concentrations in water associated with the release of gametes (Bylemans *et al.*, 2018; Bayer *et al.*, 2019; Tsuji and Shibata, 2020), spawning migration (Erickson *et al.*, 2016; Antognazza *et al.*, 2019; Thalinger *et al.*, 2019; Yatsuyanagi *et al.*, 2019), and spawning aggregation (Sakata *et al.*, 2017; Bracken *et al.*, 2019).

This thesis assesses the ability of eDNA metabarcoding to provide information on Arctic charr spawning activity hypothesising that this community-wide molecular tool can reveal species-specific peaks in read counts at the sites and times where spawning events are expected (Chapter 2).

Can eDNA metabarcoding effectively monitor priority conservation fish species in sensitive sites?

As extensively reviewed in Hering *et al.*, 2018, a set of interdependent criteria are used to evaluate the applicability of eDNA-based methods in future monitoring programmes (i.e. European Water Framework Directive - WFD) including representativeness, sensitivity and precision, comparability with other survey methods, cost-effectiveness and environmental impact. This thesis focuses on the evaluation of sensitivity and comparability of eDNA metabarcoding data to assess the efficiency of this method for detections of rare fish species (i.e. Arctic charr and coregonids) in UK lentic systems.

Certain rare fish species represent good indicators of water body deterioration given their higher sensitivity to environmental and human-induced stressors (Clarke and Murphy, 2006; Pont *et al.*, 2019), hence evaluating the capability of eDNA methods to reliably detect rare species is essential to explore the use of eDNA methods for the ecological assessment of aquatic environments. A number of studies have

demonstrated that eDNA-based surveys can be more efficient than conventional methods in detecting rare fish species in freshwater systems (Civade *et al.*, 2016; Hanfling *et al.*, 2016; Pont *et al.*, 2018; Lawson Handley *et al*., 2019; Griffiths *et al.*, 2020; Piggot *et al*., 2020) and species-specific approaches are believed to be more sensitive than eDNA metabarcoding (Lacoursière-Rousell *et al.*, 2016a; Bylemans *et al.*, 2019; Harper *et al.*, 2018).

Li *et al.*, 2018 have demonstrated that fish abundance estimates from eDNA metabarcoding are scalable allowing comparability with historical information based on conventional surveys, and suggesting eDNA metabarcoding as a good fish-based assessment tool for large-scale monitoring programmes. Nevertheless, agreement between different biomonitoring methods (targted *vs.* community-wide molecular methods and conventional *vs*. molecular methods) might be weak for species with considerable heterogeneity in abundance across sites that determines a variation in detection probability (Piggot *et al*., 2020). For low-abundant species with lower detection rates, inconsistent detections from eDNA can be confused with false positives whereas absence of detections might represent false negatives. Different methods can also have different sensitivity and affect the comparability across findings (Piggot *et al.*, 2020).

This thesis evaluates the sensitivity of eDNA metabarcoding compared to targeted qPCR for Arctic charr detections in water samples (Chapter 4), and assesses the comparability of rare species distribution and abundance data (Arctic charr and coregonids) between historical records from established methods and an existing collection of fish eDNA metabarcoding data from 101 UK freshwater bodies including sensitive sites (e.g. translocation sites; Chapter 5).

1.5 Data chapter summaries

This section provides a brief introduction to each data chapter outlining how they will address the key questions identified above.

Chapter 2: Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds

In Chapter 2 (published in *Metabarcoding & Metagenomics*) fish eDNA metabarcoding analyses were performed on water samples collected from a fishery farm holding ponds with known fish diversity and fish biomass/abundance. Fish species read counts and site occupancy from eDNA metabarcoding were hypothesised to be associated to the absolute fish biomass/abundance in each pond surveyed. In one pond, different filtration approaches and filter preservation methods were used to compare qualitative and quantitative estimates of fish eDNA metabarcoding and test consistency and repeatability of eDNA metabarcoding results between different methods of DNA capture and DNA storage.

Chapter 3: Detection of Arctic charr (*Salvelinus alpinus* L.) *spawning activity in England's biggest lake: a spatio-temporal study using environmental DNA metabarcoding*

In Chapter 3 eDNA metabarcoding was used to assess the spatial and temporal distribution of Arctic charr in Windermere north basin. Water sample collection spanned from October 2017 to July 2018, and three locations were targeted including putative and demonstrated Arctic charr spawning grounds (shore samples) and offshore sites. Peaks of Arctic charr read counts from eDNA metabarcoding were expected in autumn (Arctic charr spawning period) at the shore sites where the species migrate to spawn. Outside the spawning season, Arctic charr reads from eDNA metabarcoding were expected to be restricted to the offshore feeding grounds according to the ecology and previous eDNA observations of this deep-dwelling species (Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019). Here, we test the ability of eDNA metabarcoding to identify the spawning locations of Arctic charr in conjunction with established netting surveys.

Chapter 4: Targeted or whole-community? A "sensitive" matter for environmental DNA samples

In Chapter 4 was investigated the sensitivity of quantitative PCR and metabarcoding for detection of Arctic charr in water samples. A subset of eDNA samples (n = 23) collected in Windermere for other purposes (see Chapter 3) was used for this study. Twelve PCR replicates were performed for each molecular method and detections of Arctic charr were hypothesised to be comparable between approaches. Additionally, in this chapter was tested the hypothesis that the original concentration of vertebrate DNA and total DNA within eDNA samples influenced detection probability of the target species in PCR replicates.

Chapter 5: Environmental DNA (eDNA) metabarcoding provides accurate information of fish distribution and abundance of UK priority conservation fish

In Chapter 5, using an existing collection of fish eDNA metabarcoding data from 101 UK freshwater bodies, the distribution and relative abundance of Arctic charr, vendace and European whitefish was evaluated and compared to the expected distribution and abundance from historical data using a relative abundance scale (i.e. DAFOR scale from 0 to 5 as in Li *et al.*, 2019a). Moreover, in this chapter was tested the accuracy of eDNA metabarcoding and established methods to detect and assess the abundance of rare fish species in sensitive sites (i.e. translocation sites or site of special conservation interest). Biotic drivers of fish species composition and abiotic environmental factors were also investigated to identify significant explanatory variables determining distribution and abundance of these endangered fish species in UK freshwater systems.

Chapter 2

Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds*

The fishery ponds during drain-down. Photo credit: Dr. Watson H.V., used with permission.

Di Muri C.[1], Handley L.L.[1], Bean C.W.[2], Li J.[3,4], Peirson G.[5], Sellers G.S.[1], Walsh K.[6], Watson H.V.[1], Winfield I.J.[7] & Hänfling B.[1]

[1]Evolution and Environmental Genomics Group (@EvoHull), School of Biological and Marine Sciences, University of Hull (UoH), Cottingham Road, Hull, HU6 7RX, UK

[2]Scottish Natural Heritage, Caspian House, Mariner Court, Clydebank Business Park, Clydebank G81 2NR, UK

[3]State Key Laboratory of Marine Resources Utilization in South China Sea, Hainan University, Hainan 570228, China

[4]College of Marine Sciences, Hainan University, Haikou, Hainan 570228, China

[5]Environment Agency, Mance House, Arthur Drive, Hoo Farm Industrial Estate, Worcester Road, Kidderminster DY11 7RA, UK

[6]Environment Agency, Horizon House, Deanery Road, Bristol, BS1 5AH, UK

[7]Lake Ecosystems Group, UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK

2.1 Abstract

The sampling of environmental DNA (eDNA) coupled with cost-efficient and ever-advancing sequencing technology is propelling changes in biodiversity monitoring within aquatic ecosystems. Despite the increasing number of eDNA metabarcoding approaches, the ability to quantify species biomass and abundance in natural systems is still not fully understood. Previous studies have shown positive but sometimes weak correlations between abundance estimates from eDNA metabarcoding data and from conventional capture methods. As both methods have independent biases a lack of concordance is difficult to interpret.

Here we tested whether read counts from eDNA metabarcoding provide accurate quantitative estimates of the absolute abundance of fish in holding ponds with known fish biomass and number of individuals. The study was carried out at a fishery farm with fish ponds that were drained as part of an invasive species eradication programme. During the drain-down, all fish were counted, measured and weighed, providing absolute measures of fish species abundance and biomass for each pond of the farm.

Environmental DNA samples were collected from two ponds with the highest fish density and with the broadest species diversity found at the site. In one pond, two different DNA capture strategies (on-site filtration with enclosed filters and three different preservation buffers versus lab filtration using open filters) were used to evaluate their performance in relation to known fish community composition and biomass/abundance estimates. Fish species read counts were significantly correlated with both biomass and abundance, and this result, together with information on fish diversity, was repeatable when open or enclosed filters with different preservation buffers were used.

This research demonstrates that eDNA metabarcoding provides accurate qualitative and quantitative information on fish communities in small ponds, and results are consistent between different methods of DNA capture. The flexibility between different methods of eDNA capture and storage will be beneficial for future eDNA-based fish monitoring and their integration into fisheries management.

2.2 Introduction

Environmental DNA (eDNA) metabarcoding is revolutionising biomonitoring in aquatic environments (Harper *et al.*, 2019; Jerde, 2019; Lawson Handley, 2015; Sigsgaard *et al.*, 2019). This approach relies on the molecular identification of organisms whose genetic material has been collected, isolated and extracted from water. Species identification occurs after PCR with broad-range primers followed by High Throughput Sequencing and matching sequence reads against a reference database (see e.g. Valentini *et al.*, 2016; Deiner *et al.*, 2017 for an overview).

Environmental DNA metabarcoding has been recently suggested as a complementary biomonitoring strategy for the European Union Water Framework Directive (WFD, 2000/60/EC) which requires member states to assess the ecological status of freshwater bodies. Currently established WFD methodologies include the morphological identification and counting of phytoplankton, phytobenthos and benthic invertebrates or gillnetting and electrofishing for fish (Hering *et al.*, 2018). Yet traditional biomonitoring methods have limitations which may hamper species' detectability or correct identification. They often lack broad applicability and they frequently impact on species' welfare, such as the use of gillnets or electrofishing for fish (Radinger *et al.*, 2019). Environmental DNA metabarcoding has the advantage of detecting elusive and rare species with a lower sampling effort/cost compared to conventional methods, resolving cryptic species and identifying novel taxa through a non-invasive sampling approach (Blackman *et al.*, 2017; Grey *et al.*, 2018; Bylemans *et al.*, 2019). The ease of eDNA collection also makes this approach suitable for remote location sampling, and the molecular identification of the genetic material does not require taxonomic expertise. Environmental DNA metabarcoding has been shown to outperform established methods for the assessment of freshwater fish community composition (Civade *et al.*, 2016; Hänfling *et al.*, 2016; Valentini *et al.*, 2016; Pont *et al.*, 2018; Sard *et al.*, 2019).

The ability of eDNA metabarcoding to provide information on abundance and biomass is more controversial, and uncertainties regarding the quantitative power of eDNA metabarcoding are still present among the scientific community and monitoring agencies (Fonseca, 2018; Lamb *et al.*, 2019). This is particularly important given that

species abundance is a crucial component of biodiversity surveillance and ecological monitoring schemes, and in view of ongoing biodiversity changes worldwide (Ficetola *et al*., 2018). Positive correlations between eDNA metabarcoding data (i.e. site occupancy or read counts) and fish abundance or biomass (as deduced by established surveys e.g. gill-netting) have been demonstrated in natural environments (Thomsen *et al*., 2012; Hänfling *et al*., 2016; Lawson Handley *et al*., 2019; Sard *et al*., 2019). However, estimates from established surveys also have their own biases and may not necessarily reflect true species abundance. For example, hydroacoustics may be limited for species discrimination (Girard *et al.*, 2020), electrofishing cannot be performed in low-conductivity waters (Allard *et al*., 2014), and capture methods are usually selective and the probability of detecting rare species remains limited (MacKenzie *et al*., 2015; Pont et al., 2019). Accurate data on organism-based measures of abundance from natural aquatic habitats are difficult to obtain without exhaustive sampling – such as draining down water bodies – and hence authentic comparisons with eDNA data in natural systems are, to our knowledge, still very rare.

A second key question is how replicable eDNA metabarcoding is with different field and laboratory protocols. Standardisation of protocols may overcome this issue, but a "one-size fits all" protocol would be unrealistic (Ruppert *et al*., 2019). For instance, eDNA capture methods are often chosen based on factors such as proximity/accessibility of sampling locations and the availability of lab equipment. At present, enclosed filters are usually preferred for on-site processing, especially when remote locations are sampled, and storage buffers are used for DNA preservation within the encapsulated filter (Spens *et al*., 2017; Li *et al*., 2018; Takahashi *et al*., 2020). For field workers this approach would be logistically simple, less prone to contamination and much easier to integrate into monitoring programmes compared to laboratory-based methods of eDNA capture. Open filters allow a larger volume to be filtered, but suffer from field and transportation logistics, and are potentially more vulnerable to the risk of contamination (Li *et al*., 2018; Majaneva *et al*., 2018).

To evaluate the efficiency and suitability of different eDNA capture, a number of published studies have compared different approaches (precipitation versus filtration; on-site versus in laboratory), and a variety of filtration equipment, filters material and

filters pore size (e.g. Deiner *et al.*, 2015; Eichmiller *et al.*, 2016; Lacoursière-Roussel *et al.*, 2016; Minamoto *et al.*, 2016; Djurhuus *et al.*, 2017; Majaneva *et al.*, 2018). Recent studies have also investigated the ability of different filter types (enclosed and open filters) and preservation methods (buffers and freezing) to provide quantitative estimates of eDNA using organisms' biomass and abundance estimates from artificial stocked ponds (Li *et al.*, 2018) or from in-field visual surveys (Takahashi *et al.*, 2020). Evaluation of the quantitative performance of filter types and preservation methods based on absolute values of species biomass and abundance in natural environments would greatly contribute to the implementation of future eDNA-based surveys.

In the present study we tested whether eDNA metabarcoding can provide accurate information on the community composition and fish biomass and abundance in ponds that were drained as part of an invasive species eradication programme. During the drain down, all fish were counted, measured and weighed, providing absolute measures of species abundance and biomass, and so avoiding the biases of established techniques used in previous studies. Secondly, we tested whether estimation of fish abundance and biomass with eDNA metabarcoding is consistent between different methods of DNA capture, by comparing Sterivex (hereafter also STX) enclosed filters preserved with three different buffers (ethanol, Longmire's solution and RNAlater) and open filtration (using Mixed Cellulose Ester; MCE filters and a vacuum pump) followed by freezing preservation at -20°C.

2.3 Materials and methods

Study site and collection of fish abundance and biomass data

The study was carried out at a UK fishery farm which consisted originally of three not connected stocked ponds (Upper, Middle and Lower Lake; Fig. 1A). This site was included in an Environment Agency (EA) eradication programme for non-native topmouth gudgeon (*Pseudorasbora parva*), as part of a wider government strategy to tackle invasive species in the UK (GB Non-Native Species Secretariat, www.nonnativespecies.org). In November 2016, during the eradication programme, a new pond of 0.2 ha (hereafter "New Lake") was created, and the original three ponds drained. All fish over 150 mm total length from the original ponds were moved to the

New Lake. During relocation, fish were individually checked for potentially hidden *P. parva* individuals in their gills and mouths. The original, empty ponds were partially refilled with water and treated with the piscicide rotenone by the EA to kill all potentially remaining specimens of *P. parva*. Original ponds were left fish-free for three months (from November 2016 to January 2017). On the 18th January 2017, New Lake was completely drained and all fish were moved back to the original ponds. During fish re-allocation, individual fish were morphologically identified by experts, counted and weighed, hence the exact fish biomass and population size could be calculated for each species and water body. Following the fishery owner's request, two of the original ponds (Upper and Lower Lake) became carp ponds, and they were re-stocked mainly with *Cyprinus carpio* and a few individuals of *Perca fluviatilis* and *Carassius carassius* x *C. carpio* hybrids. Middle Lake (0.3 ha) was re-stocked with 1,248 fish with a total biomass of 634.87 kg, equivalent to 2,116.23 kg/ha. The fish community included eight species and two hybrids with biomass and number of individuals ranging from 0.7 kg/1 individual (*Squalius cephalus*) to 240.6 kg/382 individuals (*Abramis brama*) (Fig. 1B; Appendix 1: Table S1). New Lake fish community was then calculated as the sum of fish species and hybrids counted and weighed after fish re-allocation to the original ponds, and included a total number of twelve species and two hybrids with biomass and numbers ranging from 0.7-1 kg/1 individual (*S. cephalus* and *Acipenser* spp.) to 1,715.2 kg/483 individuals (*C. carpio*) (Fig. 1B; Appendix 1: Table S1). Overall, New Lake contained 2,000 fish with a total biomass of 2,695.32 kg, equivalent to 13,476.6 kg/ha. Given the diverse fish communities of New Lake and Middle Lake, our eDNA metabarcoding analyses focused on these two ponds.
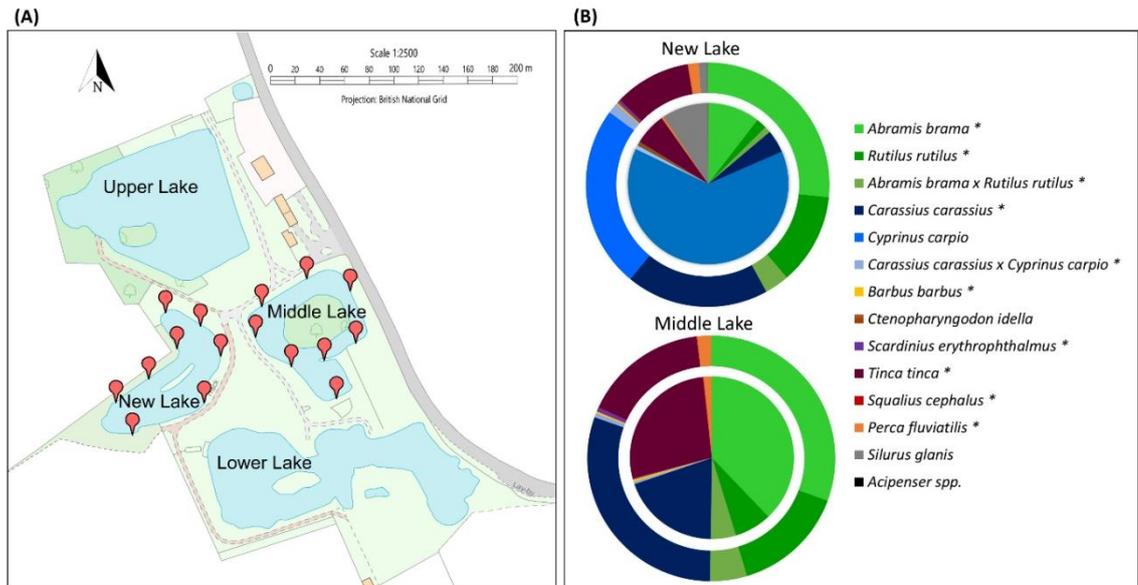
Figure 1. Map and fish diversity of the site surveyed. (**A**) Map of eDNA collection sites (in red) at the fishery venue. Map was downloaded and edited from Digimap (https://digimap.edina.ac.uk). (**B**) Fish species composition of the New Lake and Middle Lake after re-stocking (species with asterisk only). Ring pie charts (outer circles) show proportion of species composition by fish abundance (number of individuals); pie charts (inside circles) indicate proportion of species composition by fish biomass (kg).

Water sample collection, filtration, and DNA extraction

Water samples were taken on three separate occasions applying different strategies based on the goal of each occasion (see Fig. 2 for experimental design). New Lake was sampled the day before fish were transferred back to the original lakes (16th of January 2017) using MCE open filters for eDNA capture (Fig. 2). We allowed one month after reintroductions for DNA dispersion in the water, and sampled Middle Lake on the 16th and 17th of February 2017, using replicated enclosed Sterivex filters and different preservation buffers (Middle Lake-STX; Fig.2) and MCE open filters (Middle Lake-MCE; Fig. 2).
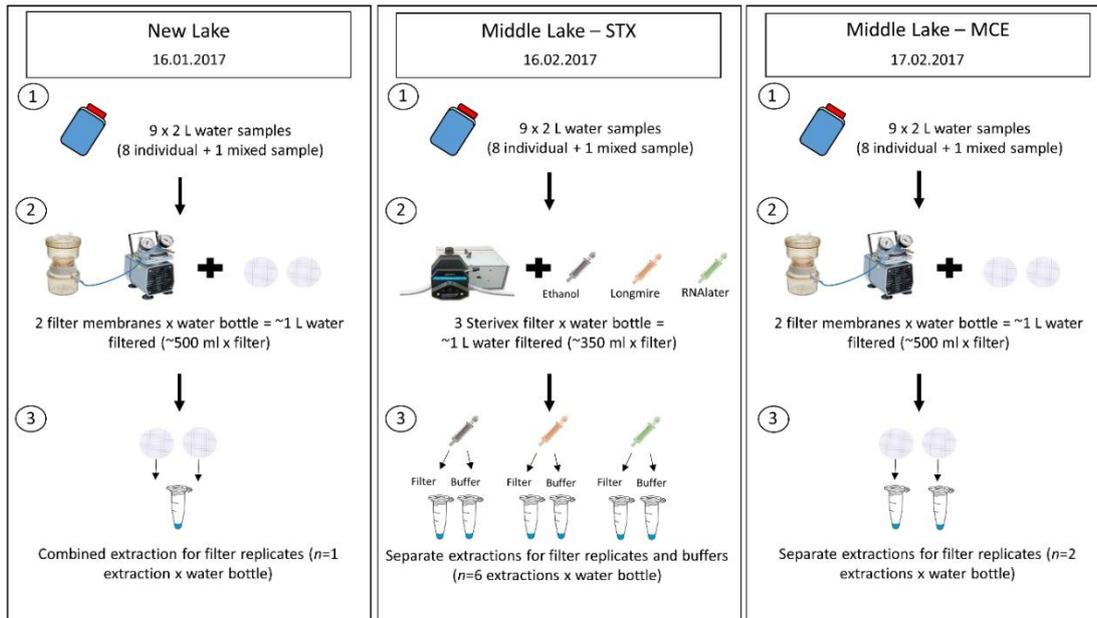
Figure 2. Experimental design. Panels show eDNA collection at different ponds (New Lake and Middle Lake) and processing strategies (Sterivex filters [STX] vs. Mixed Cellulose Ester open filters [MCE]). Numbers within the panels indicate the workflow from water sampling (1) to filtration (2) and DNA extraction (3).

All precautions to avoid contamination were taken while sampling. Fieldwork equipment was sterilised using 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK) and sterile gloves (STARLAB, UK) were changed at each sampling location. Blanks, consisting of 2 L sampling bottles filled with ultra-purified water (Milli-Q), were included for each sampling occasion. Blanks were opened once in the field and then kept and processed alongside other water samples.

On each sampling occasion, eight 2 L water samples were collected equidistantly (~30 m apart) around the perimeter of each pond (Fig. 1A). Samples were collected by hand at the water surface by pooling five 400 mL subsamples collected within a range of 5 m from the central location into a 2 L sterile plastic bottles (Gosselin™ Square HDPE, Fisher Scientific UK Ltd, UK). At each sampling occasion, immediately before filtration, a mixed sample was created using 200 mL aliquots from each of the eight water bottles collected in the field in order to evaluate differences of species detections with sampling strategies.

Samples for open filtration were placed inside cool boxes with ice packs, transported back to the laboratory and processed within 12 hours of collection. Environmental

DNA was captured on 0.45 µm MCE open filters (47 mm diameter, Whatman, GE Healthcare) using a vacuum-pump and Nalgene™ filtration units. Filtration equipment was sterilised in 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK) for 10 min, then rinsed with 5% v/v MicroSol detergent (Anachem, UK) and with purified water. Filtration was stopped after 45 min and approximately 500 mL of water was filtered through each of two MCE open filters per sample (i.e. 1 L of the 2 L total sample was filtered). MCE open filters were then stored in sterile 50 mm Petri dishes (Fisher Scientific UK Ltd, UK) sealed with parafilm (Bemis™, Fisher Scientific UK Ltd, UK) and kept at -20°C until DNA extraction.

Sterivex filtration was carried out in the field. Environmental DNA was captured using 0.45 µm Sterivex filters (PVDF membrane, Merck Millipore) connected to a peristaltic pump (Easy Load II Peristaltic Pump, In-situ Europe Ltd, UK). On-site filtration was also carried out until an individual filter became clogged, otherwise it was stopped after 45 min. Approximately 350 mL were filtered through each Sterivex filter and three Sterivex filters were used per sample. Each Sterivex filter was then preserved using 2 mL of one of three different buffers: ethanol (≥ 99.5% v/v), Longmire's solution, and RNAlater.

All DNA extractions were carried out using the mu-DNA protocol for water samples following adaptation for Sterivex as recommended in Sellers *et al*. (2018), and the DNA was eluted into 100 µL of TE buffer (see Appendix 1 for details). Replicates of MCE open filters from New Lake were co-extracted by placing both filters in a single tube for bead milling, whereas, to compare metabarcoding results of MCE open filters from the Middle Lake-MCE sampling, filter replicates were extracted separately (Fig. 2). For Sterivex filters, DNA from buffers and filters was extracted separately as previous studies have shown that DNA can become suspended in the buffer (Spens *et al*., 2017; Fig. 2). After extractions, nucleic acid yield and purity were checked on a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific).

Contamination during laboratory procedures was minimised by using separated laboratories, located on different floors, for pre-PCR and post-PCR work. Pre-PCR procedures (DNA extraction and PCR preparation) were performed in a dedicated

laboratory where only eDNA samples are handled. This laboratory has separated work stations for DNA extraction and PCR preparation. All equipment, instruments and benches are sterilised with 10% commercial bleach solution and 70% ethanol solution prior and after any work. PCR preparation occurred under UV-sterilised hoods with dedicated PCR pipettes.

Sequencing library preparation

Library preparation included a two-step PCR with a nested-tagging approach as described in Li *et al*. (2019a, 2019b). In the first round of PCR, indexed primers targeting a 106 bp region within the mitochondrial 12S gene were used (Riaz *et al*., 2011; Kelly *et al*., 2014). The first round of PCRs was performed in a final reaction volume of 25 µL including 12.5 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 1.5 µL of each indexed primer (10 µM; Integrated DNA Technologies, Belgium), 7.5 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 µL of DNA template at the original sample concentration. In order to avoid cross-contamination between samples as a consequence of evaporation/condensation and/or aerosols, reactions were prepared in 8-strip tubes with individually attached caps and covered with a drop of mineral oil (Sigma-Aldrich Company Ltd, UK). Amplifications were performed on Applied Biosystems® Veriti thermal cyclers (Life Technologies, CA, USA) with the following conditions: initial denaturation at 98°C for 5 min; 35 cycles of 98°C for 10 sec, 58°C for 20 sec and 72°C for 30 sec; final elongation step at 72°C for 7 min. Eighty-one samples, eight collection blanks, six PCR negative controls (molecular grade water), and four positive controls (genomic DNA [0.05 ng/µL] from cichlid species not occurring in the UK, *Astotilapia calliptera* and *Maylandia zebra*) were amplified in triplicate. Amplicons were checked on 2% agarose gels stained with 10,000X GelRed Nucleic Acid Gel Stain (Cambridge Bioscience, UK). Gels were imaged using Image Lab Software (Bio-Rad Laboratories Ltd, UK) to visually check for contamination in blanks/negatives, presence of target band and consistency of results among PCR replicates.

After visualisation, PCR triplicates were combined and samples belonging to the same collection site were pooled and normalised using different volumes as deduced from

strength of PCR products on gels (no/very faint band = 10 µL, faint band = 7.5 µL, bright band = 5 µL) using 1 µL of the positive samples and 5 µL of blanks/negatives for each pool (Alberdi et al., 2018).

Amplicon pools were cleaned using a double-size selection magnetic beads protocol (Bronner et al., 2013) with a ratio of 0.9X and 0.15X of magnetic beads (Mag-Bind® RXNPure Plus, Omega Bio-tek Inc, GA, USA) to PCR products (detailed protocol in Appendix 1). The double-size selection protocol and associated bead ratios were chosen to size-select the target amplicons from larger or smaller non-specific PCR products (e.g. primer dimers or non-target amplicons). Bead purification was followed by a second amplification where Illumina tags were added to each pool. Second PCRs were run in duplicate in a final reaction volume of 50 µL using 25 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 3 µL of each Illumina tag (10 µM; Integrated DNA Technologies, Belgium), 15 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 4 µL of templates. PCRs consisted of: 95°C for 3 min; 8 cycles of 98°C for 20 sec and 72°C for 1 min; and 72°C for 5 min. PCR products were checked on a 2% agarose gel alongside their non-tagged products to check for size differences after tag addition and replicates were pooled. A second double-size selection bead purification was carried out with a ratio of 0.7X and 0.15X of magnetic beads/PCR products. Tagged amplicon pools were quantified using the Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK) and pooled with equimolar concentrations into a unique library. The final library was checked for size and integrity using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA) and quantified using qPCR with the NEBNext® Library Quant Kit for Illumina® (New England Biolabs® Inc., MA, USA). The fragment size of the final library was ~ 318 bp and, following qPCR, 13 pM library was loaded on the Illumina MiSeq® with 10% PhiX using a 2 x 300 bp V3 chemistry (Illumina Inc., CA, USA).

Bioinformatics and statistical analyses

Raw sequencing data were demultiplexed using a custom Python script and subsequently analysed with metaBEAT (metaBarcoding and Environmental Analysis

Tool) v0.97.11 (https://github.com/HullUni-bioinformatics/metaBEAT), an in-house developed pipeline. Quality trimming, merging, chimera detection, clustering and taxonomic assignment against a custom-curated 12S reference database (Hänfling *et al.*, 2016) containing sequences for all UK freshwater fish species were performed. Taxonomic identity was assigned using BLAST (Zhang *et al.*, 2000) and a lowest common ancestor (LCA) approach based on the top 10% BLAST matches for any query that matched a reference sequence across more than 95% of its length at minimum identity of 100%. Unassigned sequences were subjected to a separate BLAST search against the complete NCBI nucleotide (nt) database at 100% identity. The number of reads assigned to fish species during the taxonomic assignment against the custom database (i.e. read counts) was used for downstream analyses in R v.3.5.1. (R Core Team 2018).

Total read count per sample was calculated as the sum of assigned and unassigned reads. The proportion of reads assigned to each fish species over the total read counts was then calculated on a sample by sample basis and, upon this ratio, a threshold of 0.001 (0.1%) was applied across the dataset to reduce the probability of false positives arising from cross-contamination or tag-jumping (De Barba *et al.*, 2014; Hänfling *et al.*, 2016). We only regarded a species as present in a given sample as if its ratio exceeded the threshold level. Based on the level of contamination found in sampling/filtration blanks and PCR negatives, a second arbitrary threshold was applied and all records occurring with less than 50 reads assigned were removed.

Morphological identification of fish species revealed that a substantial amount of F1 hybrids (Fig. 1; *C. carassius* x *C. carpio* and *A. brama* x *Rutilus rutilus*) were present. As community eDNA approaches are unable to differentiate hybrids from parental species these were grouped together for the purpose of our correlation analyses; i.e. data on biomass/abundance and eDNA read counts/site occupancy for hybrids and their parental species were pooled.

As read counts and site occupancy data were not normally distributed, Spearman's rank correlation coefficient was used to calculate correlations between biomass/abundance data and species average read counts and site occupancy for filter

types and treatments. Graphs were plotted using *ggplot2* (Wickham, 2016) and lines of best fit were drawn using the function *geom_smooth*. Spearman's correlation coefficients and significance levels were displayed using functions in *ggpubr* (Kassambara, 2018). Species site occupancy was calculated as the number of filter replicates with positive detections over the total number of filter replicates collected and processed using the same treatment (*n*=8).

VEGAN package v2.5-4 (Oksanen *et al.*, 2019) was then used to test differences of fish communities between filter types (Sterivex and MCE) and treatments (preservation buffers and freezing). *Betadisper* was used to investigate compositional variance of each group, and homogeneity of group dispersions was tested using ANOVA. Distances from the centroids of each treatment and the variance within treatment were visualised with a Principal Coordinates Analysis (PCoA). To test groups for compositional differences, a permutational multivariate analysis of variance (PERMANOVA), with replicates nested into each filter type, was carried out using the *adonis* function. Tests were performed on a square-root transformed abundance-weighed dissimilarity matrix (Bray-Curtis) of species composition.

Kernel density plots of fish species richness distribution across eDNA samples for each pond (New Lake and Middle Lake) and eDNA filtration/preservation strategy (Sterivex with buffers and MCE open filters replicates) were used to evaluate the number of fish species detected in the mixed samples compared to the mean species richness of eight individual samples. Density plots were built using the function *geom_density* implemented in *ggplot2* (Wickham, 2016).

Lastly, sample-based species accumulation curves (SACs) were built using the function *specaccum* for each filter type and replicate.

2.4 Results

Sequencing outputs and bioinformatics

The total number of forward and reverse sequences across 98 samples (81 eDNA samples and 17 controls) was 10,751,170. Of these, 6,398,530 paired-end sequences passed the trimming quality filter and 92% were subsequently merged. 3,389,668

sequences remained after chimera detection and clustering with an average read count per sample of 40,042 (excluding control samples). Excluding the cichlid species used as positive controls, 16 Operational Taxonomic Units (OTUs), and 1,314,623 sequences were identified as fish taxa, with thirteen OTUs remaining after applying the thresholds. All fish OTUs were identified to species level with the exceptions of records matching the family Percidae given that the 12S region amplified cannot distinguish between two species (*Perca fluviatilis* and *Sander lucioperca*) included in the database used for the taxonomic assignment. Percidae records were manually assigned to *P. fluviatilis* as this was the only species of the family identified in the study area during fish relocation.

P. parva reads found in two Middle Lake-STX samples (279 and 148 reads) were also excluded from further analyses as after eradication this species was not physically present at the site surveyed.

Environmental DNA metabarcoding fish diversity

After applying thresholds, nine OTUs of the twelve fish species translocated to New Lake remained in the final eDNA dataset. Sequences from the following taxa were detected at all eight sites within New Lake: *A. brama, C. carassius*, *C. carpio*, *P. fluviatilis, R. rutilus*, *Silurus glanis* and *Tinca tinca* (Fig. 4, Appendix 1: Fig. S1) with *C. carpio* showing the highest read counts (about 40,000) and other species reads ranging from 1,831 of *S. glanis* to 23,618 of *A. brama* (Fig. 3, Appendix 1: Fig. S1). In addition, *Barbus barbus* was detected at two sites (202 reads), and *Ctenopharyngodon idella* at one site (71 reads) (Fig. 3, 4, Appendix 1: Fig. S1). The presence of *Scardinius erythrophthalmus* was found at two sites with a low number of reads (38 and 25 reads) and, therefore, removed after applying the filter threshold (see metaBEAT raw data, Table S2 at 10.17605/OSF.IO/ZWPSQ). Taxonomic assignment based on our reference database failed to detect *Acipenser* spp., yet 79 reads (at one site) matched the family Acipenseridae during the unassigned BLAST search against GenBank, however this record was not included for further analyses as only records belonging to the taxonomic assignment against the custom fish database were considered (see unassigned blast data, Table S3 at 10.17605/OSF.IO/ZWPSQ).

All nine possible OTUs corresponding to the species reintroduced were detected beyond threshold limits in Middle Lake in both sampling occasions (16th and 17th of February). Eight OTUs (*A. brama, R. rutilus, C. carassius, C. carpio, T. tinca, B. barbus, P. fluviatilis, S. cephalus*) were detected in both Middle Lake-STX and Middle Lake-MCE, and with all filter replicates (Fig. 3, 4, Appendix 1: Fig. S1). Five of these fish OTUs (*A. brama, R. rutilus, C. carassius, C. carpio, T. tinca*) showed high site occupancy (all sites occupied) and number of reads assigned (Fig. 3, 4, Appendix 1: Fig. S1). Detection was less consistent for one of the two least abundant species, *S. erythrophthalmus*. In Middle Lake-STX, *S. erythrophthalmus* was only detected in one filter replicates preserved with RNAlater (266 reads), and in Middle Lake-MCE, in open filter replicate 1 (333 reads; Fig. 3, 4, Appendix 1: Fig. S1).

Correlation between eDNA and biomass/abundance data

We evaluated the relationship between fish eDNA read counts/site occupancy of different filter replicates and fish biomass and abundance in New Lake and Middle Lake.

We observed a strong positive association between fish read counts and fish abundance (r = 0.96; p < 0.001; Fig. 3A) and a positive, but not significant association between read counts and biomass (r = 0.75; p = 0.052; Fig. 3B) for samples collected from New Lake.

Spearman's correlations were calculated separately for each filter type (Sterivex/MCE open filters) and filter replicate for samples collected from Middle Lake (Middle Lake-STX, Middle Lake-MCE). Fish read counts for all replicates and filters were positively correlated to both fish biomass and abundance. The highest associations were observed when read counts of Sterivex filters were compared with biomass (Ethanol: r = 0.89, p = 0.019; Longmire: r = 1, p < 0.001; RNAlater: r = 0.93, p = 0.0025; Fig. 3D), and abundance (Ethanol: r = 0.89, p = 0.019; Longmire: r = 1, p < 0.001; RNAlater: r = 0.86, p = 0.014; Fig. 3C).

For MCE open filter (Middle Lake-MCE), there was a significant correlation between read counts and biomass for both filter replicates (r = 0.79, p = 0.036; r = 0.94, p =

0.0048; Fig. 3F) and between read count and abundance for MCE open filter replicate 2 (r = 0.94, p = 0.048; Fig. 3E), but the correlation between read count and abundance was not significant for MCE open filter replicate 1 (r = 0.68, p = 0.094; Fig. 3E).

A positive and significant correlation was observed when species site occupancy of New Lake eDNA samples was associated with fish abundance (r = 0.76, p = 0.049; Fig. 4A), whereas was positive but not significant the correlation with fish biomass (r = 0.58, p = 0.17; Fig. 4B).

Fish site occupancy of Middle Lake filter replicates (Middle Lake-STX, Middle Lake-MCE) was also positively correlated to both fish biomass and abundance with, however, weaker associations. Correlation coefficients and significance of the Spearman's correlations varied between filter replicates of both filter types. The strongest associations were observed when site occupancy of Sterivex filters preserved with ethanol were correlated with abundance and biomass (Ethanol: r = 0.94, p = 0.0051; Fig. 4C, 4D), but also when site occupancy of MCE open filter replicate 2 were associated with fish species biomass and abundance (r = 0.88; p = 0.021; Fig. 4E, 4F).
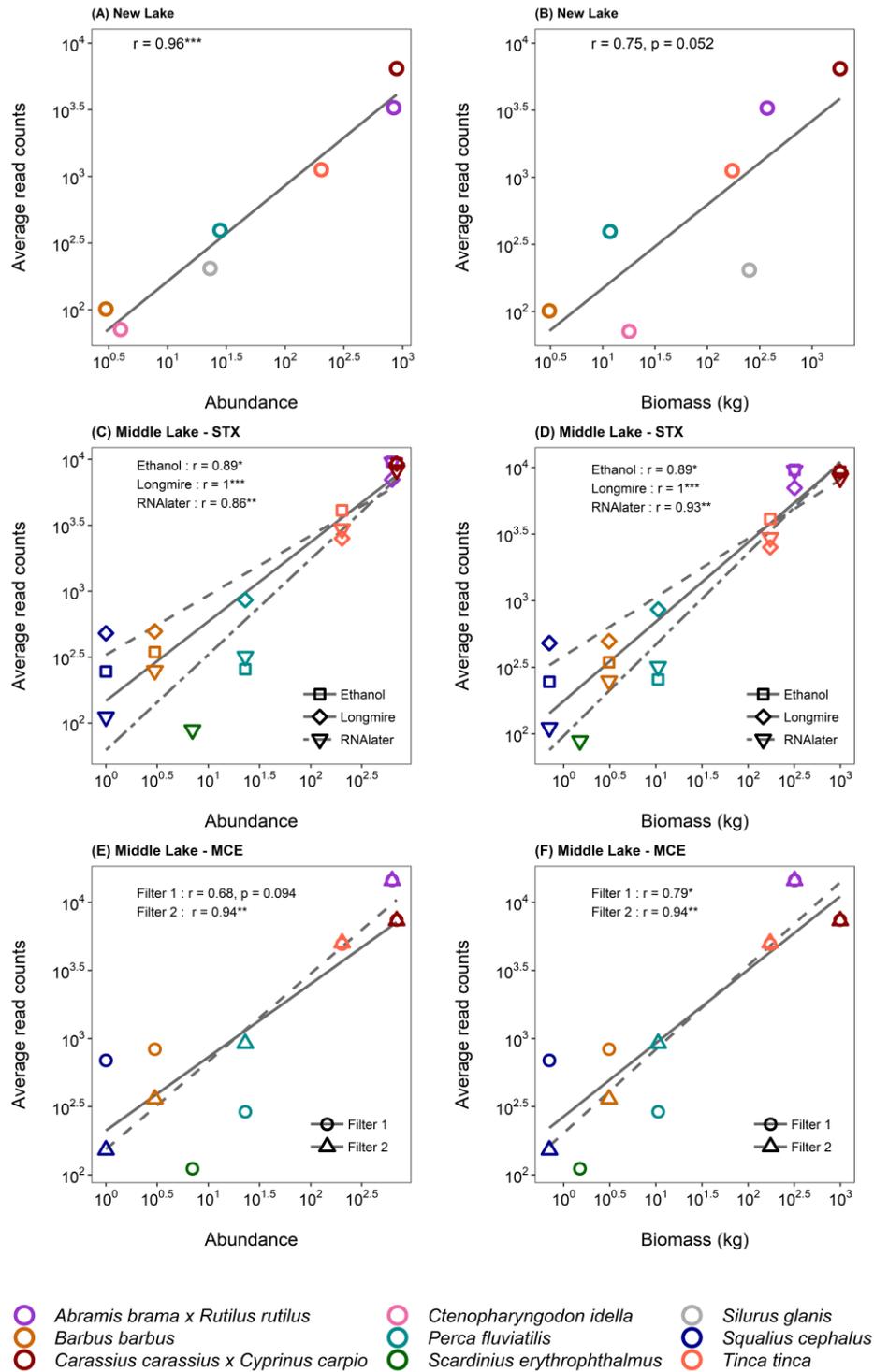
Figure 3. Correlations between eDNA metabarcoding read counts and fish abundance/biomass. Scatterplots showing lines of best fit and Spearman's correlations of fish species average read counts with abundance (number of individuals, on the left) and biomass (kg; on the right) at different sampling occasions. Panel (**A**) and (**B**) Spearman's correlations for New Lake; (**C**) and (**D**) Spearman's correlations for Middle Lake with Sterivex filters (STX); (**E**) and (**F**) Spearman's correlations for Middle Lake with open filters (MCE). Plot axes were log transformed for better visualisation. Significance codes: ***0.001; **0.01; *0.05.
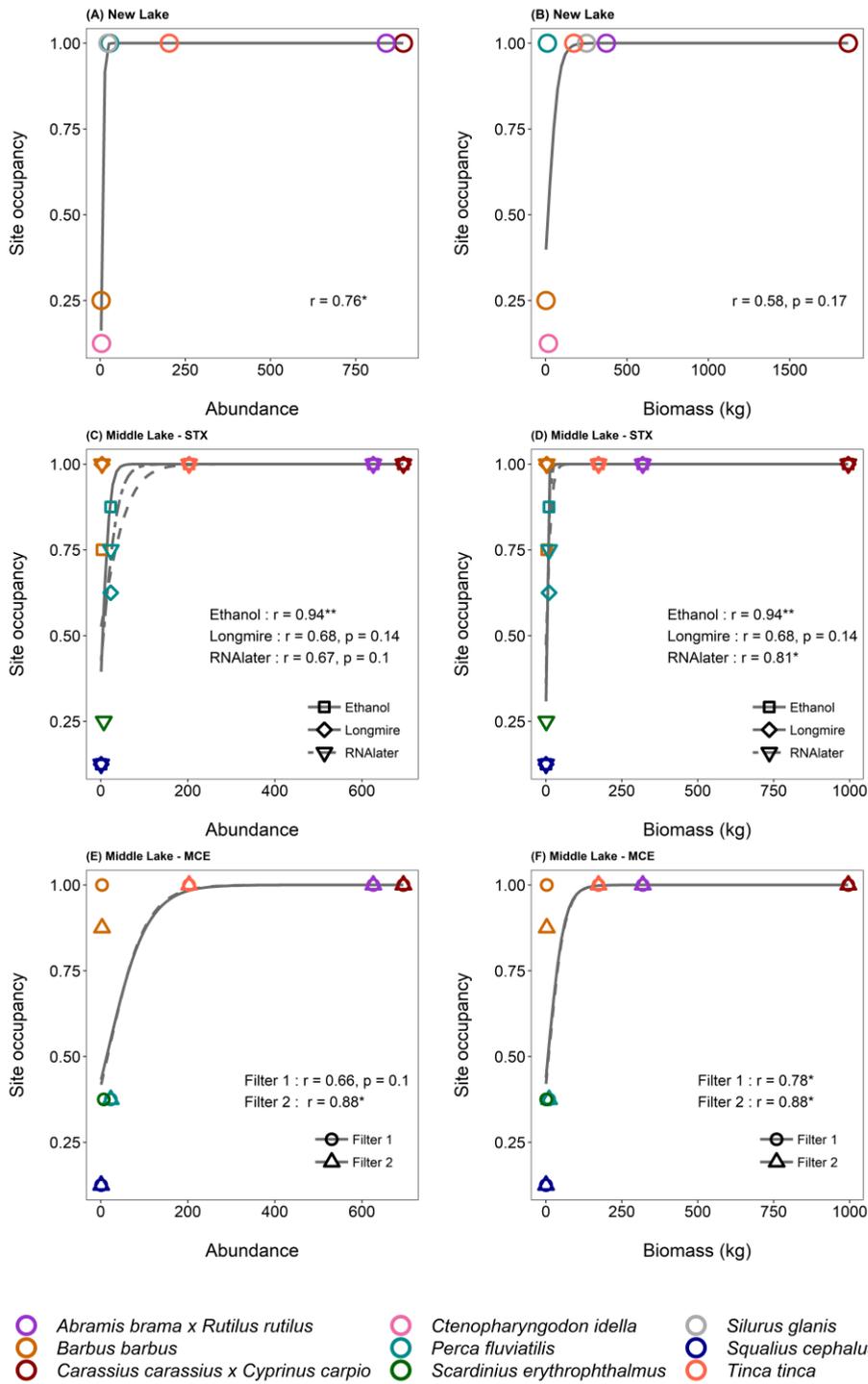
Figure 4. Correlations between eDNA metabarcoding site occupancy and fish abundance/biomass. Scatterplots showing lines of best fit and Spearman's correlations of fish species site occupancy with abundance (number of individuals, on the left) and biomass (kg; on the right) at different sampling occasions. Panel (**A**) and (**B**) Spearman's correlations for New Lake; (**C**) and (**D**) Spearman's correlations for Middle Lake with Sterivex filters (STX); (**E**) and (**F**) Spearman's correlations for Middle Lake with open filters (MCE). Significance codes: ***0.001; **0.01; *0.05. Note: mixed samples were not included in the analyses.

Effect of sampling and filtration strategies on fish community eDNA data

To evaluate the effect of different sampling strategies the mean species richness of individual samples was compared to the species richness of the mixed sample at each sampling occasion and treatment (Fig. 5A). Overall, the number of fish species detected in the mixed samples was very close, and most of the time higher, than the average number of species detected in individual field samples with the only exception of MCE filter replicate 2 (Fig. 5A). The fish species not represented in the mixed samples were often the low-occurrence taxa of the sites surveyed, and generally, excluding MCE filter 2, a number of two fish species were missing in the mixed samples. For example, in the New Lake mixed sample *B. barbus* and *C. idella* were not detected. *S. cephalus* and S. e*rythrophthalmus* were not represented in Middle Lake-STX (ethanol, RNAlater and Longmire's preservation) nor in Middle Lake-MCE filter 1 and 2 with the latter one additionally missing *B. barbus* and *P. fluviatilis*.

There were no differences between fish community composition of different filter types (ANOVA F = 0.8521, p = 0.3611; Fig. 5B) or filter replicates (ANOVA F = 0.6495, p = 0.6305; Fig. 5B).

There was no significant difference between centroids of Middle Lake fish communities described by eDNA metabarcoding when using different filter types (PERMANOVA; R^2 = 0.23278; p = 0.7231) or different preservation methods (buffers and freezing; R^2 = 0.03795; p = 0.7231). However, more variation (23%) was explained by the use of different DNA capture methods (MCE versus Sterivex), compared to within filter treatment (3.8%).

Species accumulation curves of both Sterivex and MCE filters showed that approximately six samples are required to detect all fish species in Middle Lake when filter replicates are combined (Fig. 5C). SACs of single filter replicates for Sterivex filters showed higher rates of species detection with RNAlater preservation compared to Longmire's or ethanol preservation (Fig. 5C). For the MCE open filters, most of the species were recovered with the filter replicate 1, with only a slight improvement in detection rate when the second replicate was included (Fig. 5C).
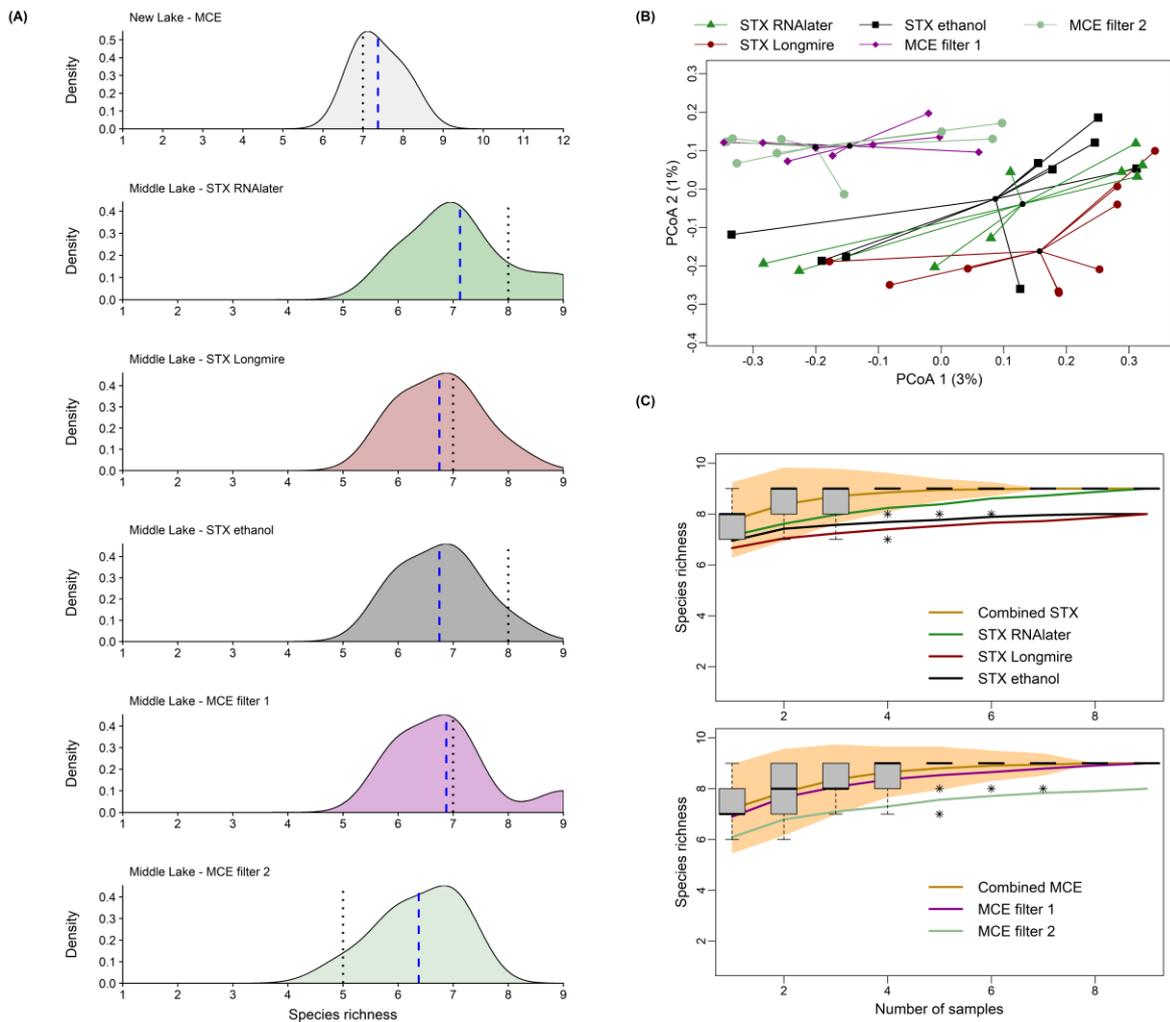
Figure 5. Environmental DNA metabarcoding fish community plots for different filter types and treatments. (**A**) Kernel density plots showing distribution of species richness across eDNA samples collected from different ponds and with different filtration and filter preservation strategies. The dashed blue lines indicate the mean species richness of individual eDNA samples (n=8), the dotted black lines indicate the species richness of mixed samples (pooled aliquots of individual samples) at each sampling occasion and filtration/preservation strategy. The x axes represent the fish species richness in each pond surveyed (New Lake = 12 species; Middle Lake = 9 species) (**B**) PCoA plot showing distances from centroids of filter types (MCE and Sterivex [STX]; ANOVA F = 0.8521, p = 0.3611) and treatments (buffers and freezing; ANOVA F = 0.6495, p = 0.6305). Distances from centroids were calculated upon a dissimilarity matrix (Bray-Curtis) of fish species read counts. (**C**) Species accumulation curves for filter replicates of Sterivex filters preserved with buffers (top) and MCE filters with freezing preservation (bottom). In both figures, golden curves are calculated based on the sum of species when filter replicates/treatments of the same filter type are combined. 95% confidence intervals refer to the golden curves and boxplots of these curves show distribution of species diversity as inferred from the method "random", which add sites in random order and was used for the SACs. Asterisks represent outliers.

2.5 Discussion

The present study demonstrates that fish read counts from eDNA metabarcoding reflect the absolute fish abundance and biomass in the ponds surveyed. Previous studies in natural environments have focused on indirect estimates of fish abundance from established surveys which have their own inherent biases. Here, we used absolute data on fish abundance and biomass from drained ponds and found that read counts from eDNA metabarcoding consistently correlate with fish abundance and often correlate with fish biomass. Moreover, the present study suggests that the use of different eDNA capture (Sterivex vs. MCE open filters) and storage methods (buffers and freezing) produce repeatable results of fish diversity, composition and biomass/abundance estimates. We additionally show that the collection of spatial and filter replicates enhances species detection probability for rare species, thus sample coverage and replication are an important consideration in experimental design.

Fish species detection

Before applying thresholds, all 12 fish taxa were successfully detected in both fishery ponds surveyed with the only exception of *S. cephalus* in New Lake (single specimen of 0.7 kg; Fig. 1B, Appendix 1: Table S1). Our findings are in line with other studies corroborating the ability of eDNA metabarcoding to describe fish diversity in lentic environments (Civade *et al*., 2016; Hänfling *et al*., 2016; Li *et al*., 2018; Lawson Handley *et al*., 2019; Zhang *et al*., 2020). The appropriate sampling effort, such as volume of water and spatial replicates collected, may vary according to the waterbody features (i.e. surface area, depth, heterogeneity) and other environmental and biological factors (Civade *et al*., 2016; de Souza *et al*., 2016; Lawson Handley *et al*., 2019). In this study, the collection of eight, evenly distributed 2 L water samples from the ponds' shore and the filtration of ~ 1 L from each water bottle provided sufficient coverage of the fish community of the ponds surveyed. An average of 1 L water filtered from each of the eight 2 L samples collected from the edge of Middle Lake was appropriate for species detection at both sampling occasions and for the different filter types used. In fact, rarefaction curves (Fig. 5C) demonstrated that when filter replicates are combined, six 2 L water samples (with ~1 L of water filtered) are sufficient to unveil the

total fish composition of this intensively stocked (0.3 ha and 2,116.23 kg/ha of fish density; Fig. 5A, C). In line with other eDNA metabarcoding studies, we suggest that near-shore sampling provides adequate species coverage as previously observed in larger and deeper lentic environments with complex fish species assemblages where a greater number of species has been detected inshore as opposed to offshore waters (Hänfling *et al.*, 2016; Zhang *et al.*, 2020). Here, we additionally highlight that an adequate sampling effort is paramount for describing species occurrence within a water body. In small, shallow lentic systems eDNA is thought to be homogeneously distributed in the water (Thomsen *et al.*, 2012; Evans *et al.*, 2017) even though the signal strength may increase closer to its source (Li *et al.*, 2019b). Yet, we demonstrated that eDNA concentration of low-abundant species DNA is very localised, hence intensive sampling efforts and the collection of an adequate number of replicates is required to detect low-occurrence taxa. For example, our mixed samples (pooled water aliquots of field individual samples) consistently detected the common fish species at all sampling occasions, but failed to detect individuals or low-abundance taxa. Spatial pooling is therefore inefficient for detailed biodiversity surveillance as suggested by Zhang *et al.* (2020), who, on a larger spatial scale with higher number of PCR replicates, still found reduced OTUs detection in mixed water samples. In line with these results, we demonstrated that eDNA detection rate is enhanced with spatial and technical replication as well as with the increased water volume filtered.

Of particular interest is the detection of *P. parva* DNA in Middle Lake samples as this invasive species was the target of the eradication programme and present in extremely high abundance before the ponds were drained and treated with a piscicide. The persistence of *P. parva* as living organisms within the pond appears extremely unlikely due to the effective eradication methods used in combination with the relatively small size of the water body (Britton *et al.*, 2008; Genovesi and Carnevali, 2011). Furthermore, *P. parva* has not been recorded in these lakes since the eradication programme. Contamination could have occurred during the water sampling or in the laboratory resulting in false positive detection (e.g. Hänfling *et al.* 2016). However, no *P. parva* DNA was found in any of the control samples or in the water samples from New Lake. A possible explanation for this record is that *P. parva* eDNA originated from sediment re-suspension in the water column during sampling or from carcasses

remaining at the bottom of the pond. Environmental DNA is known to be less concentrated and less persistent in water compared to sediment, where it remains detectable for over three months also when species are removed from the system (Turner *et al.*, 2015). A further reasonable option would be to consider this result as a true record even if we have no evidence that the species re-colonised the pond after the eradication. Previous studies have suggested that *P. parva* may suffer from recruitment failure and local extirpation when population numbers are low due to human or natural disturbance (Copp *et al.*, 2007; Davidson *et al.*, 2017). Therefore, when monitoring the success of eradication attempts, extreme caution should be taken with false-positive or false-negative detections for the target species and the use of conventional methods to corroborate eDNA detections has been recommended (Davison *et al.*, 2019; Robinson *et al.*, 2019).

Read counts correlate to biomass and abundance

Our study investigated the correlation between eDNA metabarcoding data and actual measures of species biomass and abundance in semi-natural lentic systems. Takahara *et al.* (2012) previously demonstrated that eDNA concentrations in ponds artificially stocked with *C. carpio* reflected the species absolute biomass. Using a similar environmental setting, we found that our eDNA metabarcoding results accurately reflect abundance patterns and reveal positive and mostly significant correlations between read counts and fish species biomass (weight) and abundance (Fig. 3). Recently Kelly *et al.* (2019) demonstrated that when amplification efficiency is high in PCR-based studies, proportional indices of eDNA reads capture trends in taxon biomass with high accuracy. Our study supports these results as we found that the species' read counts were an accurate quantitative parameter to describe taxon biomass and abundance.

Positive associations were observed between species site occupancy and fish biomass/abundance, however, less significant than correlations with read counts (Fig. 4). In our study system, the relatively small size of the water bodies surveyed, coupled with the high fish densities, resulted in relatively homogeneous distribution of the common species' eDNA, generating a better representation of fish biomass and

abundance when read counts were used for quantitative inferences. In larger and heterogeneous lentic environments, the spatial variation of the species' eDNA signal is likely to be as or more reliable than read counts for quantitative estimates (Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019; Sard *et al.*, 2019).

Current uncertainties regarding the quantitative power of eDNA metabarcoding ultimately originate from our lack of knowledge on the origin and fate of eDNA in aquatic systems (Klymus *et al.*, 2015; Lacoursière-Roussel *et al.*, 2016; Sassoubre *et al.*, 2016). Age, physiology, life history and metabolic rate all play a role in the amount of eDNA released (eDNA shedding rate) from organisms into their surroundings (Barnes *et al.*, 2014; Goldberg *et al.*, 2016; Ruppert *et al.*, 2019). Physical, chemical and biological forces such as dilution, sedimentation and re-suspension, hydrolysis, oxidation and microbial activity, can all influence eDNA persistence and dynamics within aquatic habitats (Turner *et al.*, 2015). In addition, the degradation of genetic material is also promoted by high temperature and acidity (Seymour *et al.*, 2018; Ruppert *et al.*, 2019). In our study system, the fish age distribution was relatively narrow (personal observation), therefore, reducing the effect of different eDNA shedding rates from distinct life stages and age classes. Moreover, the ponds surveyed were similar in terms of high fish density and species composition and were also exposed to stable environmental conditions that positively influenced the reproducible and reliable quantitative characterisation of the fish communities investigated.

A lack of robust sampling and metabarcoding protocols may also contribute to a distortion of the observed diversity patterns. Insufficient sampling effort, inhibition, primer biases, sequencing artefacts, database inaccuracy and contamination are the main methodological sources of bias (McKee *et al.*, 2015; Grey *et al.*, 2018; Collins *et al.*, 2019; Wood *et al.*, 2019). In the present study, the quantitative fish assessment of the two ponds surveyed demonstrates the accuracy of optimised eDNA metabarcoding protocols to reflect species biomass and abundance. In recent years, sampling, laboratory, and bioinformatics workflows have been progressively refined for the characterisation of fish communities within UK freshwater ecosystems (Hänfling *et al.*, 2016; Li *et al.*, 2018; Lawson Handley *et al.*, 2019; Li *et al.*, 2019a). Here, we have demonstrated that optimised sampling strategy, enhanced extraction protocol with an

additional inhibitor removal step (Sellers *et al.*, 2018), replication during PCRs and development of a custom-curated database with new reference sequences, strengthened the probability of detection, reduce taxonomic assignment bias, and overall provided reliable quantitative data of fish biomass and abundance.

Suitable eDNA metabarcoding data for quantitative fish monitoring require comparable measures of biomass and abundance across studies and over time, for example, to detect trends in abundance of fish populations. In light of this, the use of site occupancy appears a more practical approach as abundance/biomass estimates from site occupancy are easily comparable even across studies with uneven sampling efforts (Li *et al.*, 2019a). Conversely, quantitative estimates from eDNA metabarcoding read counts will need to be adjusted to standardised metrics. One possible approach, proposed by Ushio *et al.* (2018), is the use of internal standards to convert raw sequence reads from metabarcoding into standardised reads using standard curves of known copy number.

Impact of DNA capture and preservation methods

In our study the correlations between sequence read counts and species abundance/biomass were consistently high for all filtration treatments with average correlation coefficients of 0.93 for Sterivex filters and 0.84 for MCE filters (Fig. 3). The variation of correlations observed between filter types may be explained by differences between read counts assigned to species as a result of different water volumes filtered between Sterivex and MCE filters (Appendix 1: Fig. S1). However, for equal amounts of water filtered and high DNA concentrations, MCE open filters usually capture a higher amount of DNA compared to enclosed Sterivex filters possibly due to the tendency of Sterivex units to clog more easily (Li *et al.*, 2018; Takahashi *et al.*, 2020). Quantitative differences between filter types may also vary with the target species as observed in this study (Takahashi *et al.*, 2020). In fact, while we observed a general trend of higher species read counts in MCE filters, we also observed the opposite trend for *C. carpio* which showed lower reads in MCE filter replicates compared to Sterivex filters (Appendix 1: Fig. S1).

The higher species richness found in Sterivex filters preserved with RNAlater and MCE open filter replicate 1 resulted from the detection of only one low-abundant taxon within the pond (*S. erythropthalmus*; Fig. 1). We therefore consider this result a stochastic effect between filter replicates or storage methods.

Overall, we found that both filter types showed a good representation of fish diversity and community composition and, consequently, we suggest that they can be used interchangeably depending on time, resources and location of the study. Sterivex filters, for instance, are effective for field processing of water samples, facilitating collection in remote locations. After sample collection, Sterivex are immediately filtered on-site (using peristaltic pumps or sterile syringes) and the risk of contamination is reduced because of the lack of filter handling (Spens *et al.*, 2017; Li *et al.*, 2018). In the present study, there was no evidence of higher contamination in MCE open filters compared to Sterivex filters indicating that preventing on-site and in-lab contaminations is sufficient to minimise/avoid DNA contaminations regardless of the filter types' choice. The use of Sterivex filters, or enclosed filters in general, is however more amenable to large-scale monitoring programmes for environmental managers or citizen science projects (Biggs *et al.*, 2015; Buxton *et al.*, 2018; Larson *et al.*, 2020). Nevertheless, Sterivex filters are currently almost 15 times more expensive than MCE open filters, DNA extraction is more time-consuming, and, when syringes are used for filtration, the Sterivex method requires a large amount of disposable plastic consumables. The use of pre-packed sterile syringes is nonetheless preferred over pumps' suction (vacuum or peristaltic) to reduce filtration time (personal observation; Li *et al.*, 2018).

2.6 Conclusion

This study underpins valuable considerations for the quantitative estimates of eDNA metabarcoding data. We demonstrated that eDNA metabarcoding data correlate with actual abundance and biomass of fish communities within small freshwater systems with high fish density.

Established methods (i.e. hydroacoustic, electrofishing, gillnetting) for obtaining quantitative estimates of fish abundance are resource intensive and may not be

suitable for all water bodies and species (Winfield *et al.*, 2009). Furthermore, quantitative interpretation of data is often complex (hydroacoustic) or relies on large sampling effort (netting/electrofishing) (Winfield *et al.*, 2009), hence becoming costly in terms of financial, human resources and habitat disturbance or species mortality. More importantly, these methods can be prone to errors as they are not exhaustive sampling methods and, therefore, can only provide approximation of species abundance.

Environmental DNA metabarcoding is arguably a more flexible tool, adaptable to all aquatic environments and fish species, is non-lethal, and the sources of errors can be minimised through a careful optimisation of field and laboratory protocols.

Monitoring trends in population size and community structure is paramount to the assessment of species health and viability, and the outputs are required to undertake management actions and to guide conservation decisions (Kull *et al.*, 2008). Implementation of eDNA metabarcoding will drive a step-change towards non-invasive monitoring strategies for next-generation ecosystems surveillance. Environmental DNA metabarcoding, as a non-invasive, fast, universally applicable approach, is nowadays claiming the attention of researchers, stakeholders and governmental agencies. Therefore, exploring, evaluating and finally establishing the quantitative value of such a broadly-used tool for diversity monitoring is essential.

Funding

Acknowledgements

Shum, Jairo Arroyave and Stephanie McLean for the constructive feedback on the manuscript before its initial submission.

Data Resources

Chapter 3

Detection of Arctic charr (*Salvelinus alpinus* L.) spawning activity in England's biggest lake: a spatio-temporal study using environmental DNA metabarcoding

A day of sampling at Lake Windermere (Cumbria, UK) with B.H. and H.V.W.
Photo credit: Dr. Winfield I.J., used with permission.

3.1 Abstract

Outlining the timing and location of fish reproductive events is crucial for the implementation of correct management and conservation schemes. Conventional methods to monitor these events are often unable to assess the spawning effort or can be invasive and therefore problematic, especially when threatened fish populations are the target, such as the Arctic charr (*Salvelinus alpinus* L.) populations in Windermere (Cumbria, UK). Arctic charr populations have been studied in this lake since 1960, and locations and characteristics of spawning grounds have been described in detail using traditional techniques such as hydroacoustic, physical and visual surveys of the lake bottom. Here, we used environmental DNA (eDNA) metabarcoding to

assess the spatial distribution of the Arctic charr throughout the year to test whether this technique could allow us to identify spawning locations and activity in conjunction with traditional netting surveys. Seasonal sampling was carried out between October 2017 and July 2018 at three locations in the lake, covering putative and demonstrated spawning sites. In this study, eDNA metabarcoding enabled accurate spatial and temporal characterisation of Arctic charr spawning events as well as a description of fish community changes at the sampling sites. Peaks of Arctic charr read counts from eDNA metabarcoding were observed during the spawning season and at specific locations of both putative and demonstrated spawning sites. The catches of Arctic charr mature individuals confirmed the association between the Arctic charr spawning activity and the peaks of eDNA metabarcoding read counts. Our observations revealed the potential of eDNA metabarcoding to characterise the spawning of the Arctic charr and to reveal fish species interactions that can be beneficial for the long-term monitoring of Arctic charr populations in the lake.

3.2 Introduction

Anthropogenic pressures are threatening freshwater fish populations worldwide, and conservation biologists and environmental managers are striving to preserve such diversity as it provides ecosystem services to humans and it holds an intrinsic evolutionary and ecological value (IUCN FFSG, 2015; Piccolo, 2015; Lynch *et al.*, 2016). A number of emblematic species of conservation interest belong to the family Salmonidae, and many salmonid populations have declined dramatically in recent decades mirroring a general loss of biodiversity (Sterneker, Denic and Geist, 2014; Piccolo, 2015). The Arctic charr, *Salvelinus alpinus* (Linnaeus, 1758) is a salmonid species and it is the northernmost of all anadromous and freshwater fish on Earth (Hansen *et al.*, 2019) with a distribution ranging from the southernmost temperate areas of the eastern North America and the Alps in Europe, to the northernmost strip of land in Eurasia and North America (Johnson, 1980). At the dawn of the most recent post-glacial era, the subsequent isolation of many Arctic charr populations in post-glacial lakes has led to high polymorphism and plasticity among and within local populations (Klemetsen, 2010; Jonsson and Jonsson, 2001). Such polymorphism encompasses life-history tactics (e.g. anadromous and non-migratory forms;

Klemetsen *et al.*, 2003), ecophenotypes (specialisation in diet and habitat preferences associated to morphological variation; Adams *et al.*, 2003; Klemetsen *et al.*, 2006; Power *et al.*, 2009), and reproductive strategies (Frost, 1965; Telnes and Saegrov, 2004; Smalås, Amundsen and Knudsen, 2013). Many of such variations are not mutually exclusive and have a genetic basis (Gomez-Uchida *et al.*, 2008; Power *et al.*, 2009; Corrigan *et al.*, 2011).

This high diversity reflects the ability of Arctic charr populations to adapt to cold, nutrient impoverished and species-poor lacustrine habitats with a variety of niches available (Jonsson and Jonsson, 2001; Reist *et al.*, 2013). However, the ongoing anthropogenic changes are determining a contraction of the species' distribution and many populations have gone extinct (Freyhof and Kottelat, 2005; Maitland *et al.*, 2007; Alekseyev *et al.*, 2013). The conservation status of Arctic charr varies throughout its distribution, but generally the level of protection is low and, for example, this species does not appear in the Red Data Book and it is not listed in the European Union Habitats and Species Directive (Adams *et al.*, 2007).

In the UK, the Arctic charr is a priority conservation taxon with a Biodiversity Action Plan because of the limited distribution, the past extirpations and the concerns over the conservation status of many populations (Maitland *et al.,* 2007). For example, in Scotland 295 lochs were known to support this species, however at least 12 populations are now considered extinct and many others have not been monitored in recent years (Maitland *et al.*, 2007; Adams and Maitland, 2018). The distribution throughout the remainder of the UK is limited to eight populations in the English Lake District (Cumbria) and four populations in Wales (Maitland *et al.*, 2007). A number of human-driven changes are causing the decline of Arctic charr populations in the UK, and many of such threats interfere with the species reproductive activity, and hence with a correct recruitment that is essential to maintain self-sustaining populations (Adams *et al.*, 2007).

The Arctic charr is an iteroparous species, typically spawning several times during its lifetime and, similarly to other salmonids, the Arctic charr is a lithophilic fish and mature adults release their gametes at shallow gravel banks where the females dig

depressions, named redds, for eggs incubation (Groot, 1996; Esteve, 2005). Here, on shallow waters, mature females attract spawning males which court females or compete with other males trying to fertilise the eggs (Esteve, 2005; Brattli *et al.*, 2018). For successful reproduction, all salmonids, including the Arctic charr, depend on clean substrate (low-fine sediment and well-oxygenated interstitial zone) for the correct eggs and larval development (Sterneker, Denic and Geist, 2014). A number of studies recognised siltation and sedimentation of spawning gravels as major factors for reproductive failure (Levasseur *et al.*, 2006; Franssen *et al.*, 2012). Moreover, the global warming, decreasing oxygen solubility in water, is expected to further reduce the recruitment success of such species (Reist *et al.*, 2006; Winfield *et al.*, 2010; Kelly *et al.*, 2020).

Annually, breeding Arctic charr gather on specific shallow spawning grounds, and such spawning areas in Lake Windermere (Cumbria, Lake District National Park, UK), have been accurately described over the last 50 years (Frost, 1965; Miller *et al.*, 2015). Arctic charr is a species of prominent interest in Windermere for its long-standing economical, ecological and cultural value (Winfield *et al.*, 2008a, Winfield *et al.*, 2019), and historically, the Arctic charr has been the most abundant fish in the lake (Frost, 1977; Mills, 1990). However, eutrophication, climate change, and competition with introduced species (i.e. roach, *Rutilus rutilus*) are driving the decline of Arctic charr populations in this lake (Winfield *et al.*, 2008a; Winfield *et al.*, 2015; Winfield *et al.*, 2019).

In 1965, pioneering research on breeding habitats of the Arctic charr in Windermere described two sympatric populations in the mesotrophic north basin and the eutrophic south basin, with autumn and spring spawning events, and their genetic and phenotypic divergence (Frost, 1965; Corrigan *et al.*, 2011). Autumn-spawners release their gametes at depths of around 2 m between November and December, whereas spring-spawners mature between February and March and spawn at deeper sites between 15 and 20 m. Both the north and south basins in Windermere sustain autumn and spring-spawners, and a variety of putative and demonstrated spawning locations have been described in the lake (Frost 1965; Miller et al., 2015; Winfield *et al.*, 2015).

Changes in the conditions of spawning grounds could contribute to the observed Arctic charr decline in Windermere, therefore Miller *et al.*, 2015 and Winfield *et al.*, 2015 carried out a recent assessment of these key sites. Hydroacoustic, physical and visual surveys of the lake bottom combined with historical and contemporary netting data were used to evaluate the present suitability of putative and demonstrated Arctic charr spawning grounds described decades before (Miller *et al.*, 2015; Winfield *et al.*, 2015). These observations suggested a general deterioration of the sites surveyed due to increased sedimentation levels, which were particularly pronounced within the south basin where only 12% of the monitored sites showed optimal or sub-optimal conditions (Miller *et al.*, 2015). Moreover, hydroacoustic surveys indicated that these sites, originally described as spawning places, may in fact be aggregation areas, and the act of spawning is possibly confined to specific locations within these aggregation areas and limited to depths below 5 m where hard substrates still occur (Miller *et al.*, 2015, Winfield *et al.*, 2015). Limitation of hard substrates to the shallowest areas of the lake pose important conservation concerns for spring-spawning Arctic charr populations with deeper breeding grounds (Winfield *et al.*, 2008a).

Extensive monitoring of spawning activity at depths less than 5 m is challenging using established survey methods. Hydroacoustic applications can be limited in shallow inshore areas (Miller *et al.*, 2015), and netting surveys cannot be deployed widely due to ethical implications, especially when species of conservation interest are the target. Novel, non-invasive and broadly applicable monitoring approaches are therefore required to characterise times and locations of spawning activity. The use of molecular approaches applied to the analysis of environmental DNA (eDNA), has recently transformed the fields of conservation and resource management (Harper *et al.*, 2019; Jerde 2019; Holman *et al.*, 2019). Environmental DNA is the genetic material shed by organisms into their surroundings. This genetic footprint can be captured from the environment (i.e. in water samples) and identified using targeted or whole-community molecular approaches (see Lawson Handley *et al.*, 2015 or Rees *et al.*, 2014 for an overview). Targeted approaches can be species-specific and rely on the amplification of unique DNA sequences within the environmental matrix. The whole-community approach, named metabarcoding, can accurately describe the diversity of a certain taxonomic group down to haplotype identification (Ruppert *et al.*, 2019; Deiner *et al.*,

2017; Elbrecht *et al.*, 2018; Marshall and Stepien 2019). The parallel amplification of informative DNA regions is achieved by using conserved, broad-range primers and high-throughput sequencing. Sequences generated are compared to reference sequences for taxonomic assignment. Environmental DNA metabarcoding has been used for qualitative and semi-quantitative surveys of various aquatic environments, from ponds and lakes (Harper *et al.*, 2019; Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019), to rivers, estuaries and marine systems (Pont *et al.*, 2019; Sigsgaard *et al.*, 2019; Stoeckle *et al.*, 2017a).

Species eDNA concentrations in aquatic systems are known to fluctuate according to several abiotic and biotic factors, including organisms' reproductive activity (see Stewart 2019 for an overview). Temporally and spatially constrained changes in species eDNA concentration have been used as indicators of reproductive migrations and/or aggregation of several aquatic species: bigheaded carp (*Hypophthalmichthys nobilis and H. molitrix;* Erickson *et al.*, 2016), sockeye salmon (*Oncorhynchus nerka*; Tillotson *et al.*, 2018), European shad (*Alosa* spp.; Antognazza *et al.*, 2019), Danube bleak and Vimba bream (*Alburnus mento* and *Vimba vimba*; Thalinger *et al.*, 2019), Shishamo smelt (*Spirinchus lanceolatus*; Yatsuyanagi *et al.*, 2020), and sea lamprey (*Petromyzon marinus*; Bracken *et al.*, 2019). Species-specific peaks of eDNA concentration have been also associated with the release of gametes by aquatic organisms during external fertilisation. For instance, Bylemans *et al.* (2017) considered sperm as the primary source of eDNA and accounted for their genomic features (high nuclear DNA/low mitochondrial DNA) to identify altered ratios of mitochondrial DNA to nuclear DNA of Macquarie perch (*Macquaria australasica*) as a sign of a spawning event. These studies indicate the ability of targeted eDNA analysis to reveal reproductive features of aquatic organisms.

The capacity of eDNA metabarcoding to produce quantitative estimates has also been harnessed, where studies have demonstrated the correlation between read counts and relative abundance/biomass for different species (Chapter 2; Hänfling *et al.*, 2016; Li *et al.*, 2018). Based on evidence provided by previous eDNA research, we hypothesise (1) that eDNA metabarcoding analyses, using extensive spatio-temporal water sampling, can detect aquatic species spawning activity as an upshot of the temporal and spatial variation in eDNA concentrations during the breeding season, (2)

that species-specific peaks in read counts from eDNA metabarcoding reflect the sites and times where spawning events are expected, and (3) that the fish community data from eDNA metabarcoding can be used to monitor fish species assemblages and predict future changes of Arctic charr populations.

Within this study, we focused on the shallowest breeding grounds of the autumn-spawning Arctic charr population in Windermere's north basin that have been assessed as suitable to support spawning activity (Miller *et al.*, 2015; Winfield *et al.*, 2015). Specifically, we targeted putative and demonstrated spawning grounds, with the latter ones monitored yearly by the Centre of Ecology and Hydrology (CEH, Lancaster, UK) and the Freshwater Biological Association (FBA, Windermere outpost, UK) for assessing trends of the autumn-spawning Arctic charr population within the lake. Here, results of the gill-netting surveys were used to confirm that peaks in Arctic charr read counts were associated with spawning individuals caught at these sites. Additionally, in autumn, we expected an eDNA signal of similar strength at the putative spawning grounds if any spawning activity was occurring.

3.3 Materials and methods

Study site

Water sampling was carried out in Lake Windermere (UK) at the known spawning site of North Thompson Holme island (SNTH, three locations), and along two transects located at the west shore of Red Nab (SRN, eight locations) and at offshore locations approximately in the middle of the lake (OF, five locations) reflecting further putative spawning locations and deep water feeding habitats respectively (Fig. 1). At SNTH, catches of mature Arctic charr individuals from twelve gill-netting surveys carried out between October and December 2017 were used to verify whether spawning activity had taken place at the time of eDNA sampling. At SRN, Arctic charr spawning activity has not been monitored or demonstrated through catches of spawning individuals. However, the area has been identified as a putative spawning ground based on anecdotal historical records, presence of suitable substrate, and Passive Acoustic Monitoring (PAM) data which identified noises connected to Arctic charr spawning

activities (gravel displacement or sounds associated with air exchange with swimbladder regulation; Bolgan *et al.*, 2017).
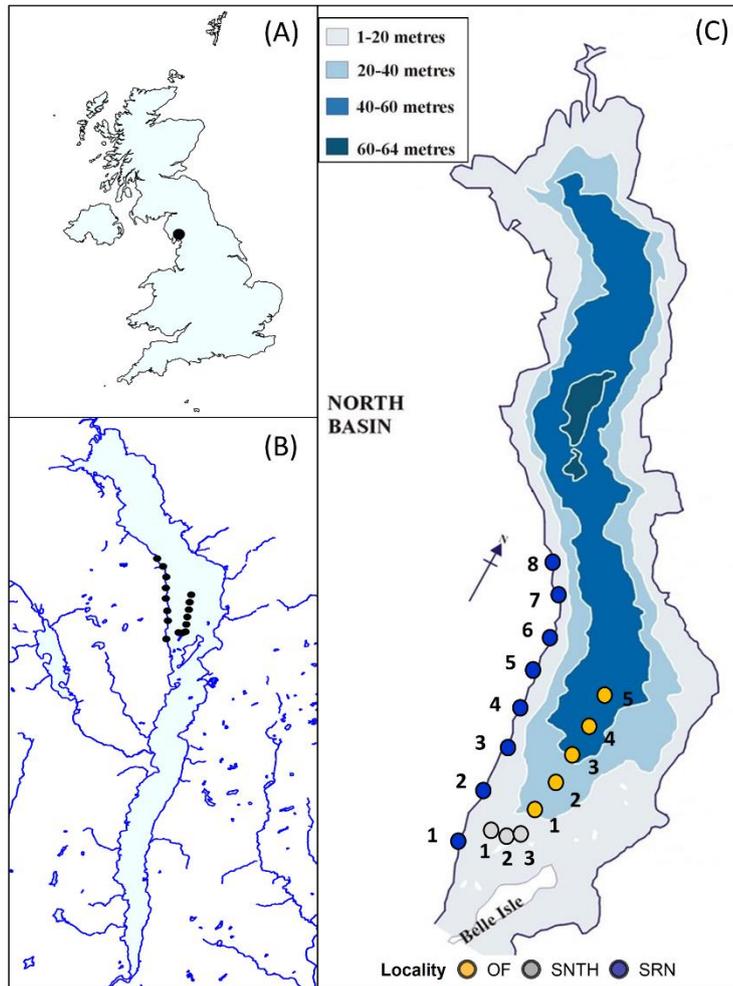
Figure 1. Map of Windermere and eDNA collection sites. (**A**) Lake Windermere's location in the Lake District, Cumbria, UK; (**B**) Location of eDNA sampling sites in Lake Windermere's north basin; (**C**) detailed bathymetric map of Windermere's north basin with sites and localities sampled during our eDNA surveys where "OF" are offshore sites (Arctic charr feeding grounds), "SNTH" are inshore sites located at the shore of North Thompson Holme island (Arctic charr monitored spawning grounds) and "SRN" are shoreline sites on the west side of the lake (Arctic charr putative spawning grounds). Edited from Ramsbottom (1976) and used with permission of the Freshwater Biological Association.

Water sample collection, filtration, and DNA extraction

Water samples were collected during twelve different dates between October 2017 and July 2018, with a higher sampling effort in autumn when Arctic charr are expected to be reproductively active (Fig. 2). Given the inability to collect water samples from

some locations on some dates due to logistic reasons (i.e. boat not available), results presented in the main text focus on completed collection events (Fig. 2). Arctic charr detections from additional eDNA sampling dates can be found in the Appendix 2 (Fig. S1).

At each site in the onshore location (SRN), five spatial replicates (5 x 400 mL taken across 50 m) merged into a single 2 L sterile plastic bottle (Gosselin™ Square HDPE, Fisher Scientific UK Ltd, UK) were collected at the surface water layer. One sampling blank, consisting of a 2 L sterile bottle filled with ultra-purified water (Milli-Q), was used and opened once in the field. In the offshore sites, SNTH and OF, samples were collected at different depths (2 m to 40 m; Fig. 1) from a boat using a 1.5 L, Friedinger-like water sampler. The water sampler was lowered three times at each sampling site (3 x ~650 mL) in order to collect spatial replicates, subsequently merged into a single 2 L sterile plastic bottle. Two sampling blanks were used during the offshore sample collection at the beginning and end of the water sampling to account for contaminations introduced by the use of a water sampler. After bleaching, the Milli-Q water of each blank was used to rinse the water sampler before pouring it back into the 2 L bottles.

To minimise cross-contamination between sites, nitrile sterile gloves (STARLAB, UK) were worn all time and changed between collection sites. The water sampler was sterilised, while moving between sites, using 10 min soaking in a 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK) followed by rinsing with 5% v/v MicroSol detergent (Anachem, UK) and purified water. At each site, the sampler was also quickly lowered and washed with the lake's water before collection occurred. Site coordinates were recorded during the first eDNA sampling event (20 October 2017) using a hand held Geographic Positioning System (GPS) (Garmin eTrex 10, Kansas, USA; Appendix 2: Table S1), and these coordinates were used to navigate to the collection sites during all subsequent sampling events.

Water bottles were kept in cool boxes covered with ice packs and filtered within six hours maximum after water collection. Water was filtered using vacuum-pumps coupled with Nalgene™ units and DNA was captured onto 0.45 μm mixed cellulose

ester filters (47 mm diameter, Whatman, GE Healthcare). Generally, two filters were used for 2 L of water collected from the shoreline sites and one filter was used for water samples collected offshore with a few exceptions. A filtration blank was run during each filtration round, where 2 L of ultra-purified water were filtered alongside water samples and sampling blanks. Filters were stored in sterile 50 mm Petri dishes (Fisher Scientific UK Ltd, UK) sealed with parafilm (Bemis™, Fisher Scientific UK Ltd, UK), and kept at -20° C until extraction.

The mu-DNA water protocol (Sellers *et al*. 2018) was used for DNA extraction from filters, where samples, sampling blanks, and filtration blanks belonging to different sampling dates were extracted separately. An extraction blank, consisting only of extraction reagents, was included for each extraction round. When reagents were newly prepared and used for extractions, DNA yield and purity was checked using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific).

Sequencing library preparation

The sequencing library was built using a custom library preparation protocol with double-tagging (Li *et al*., 2019). For the first PCR, indexed primers amplifying a ~106 bp region of the mitochondrial 12S ribosomal RNA (rRNA) in fish were used (Kelly *et al*. 2014; Riaz *et al*. 2011). PCR was performed with a final reaction volume of 25 µL, including 12.5 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 1.5 µL of each indexed primer (10 µM; Integrated DNA Technologies, Belgium), 0.5 µL of the Thermo Scientific™ Bovine Serum Albumin (Fisher Scientific UK Ltd, UK), 7 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 µL of DNA template at the original sample concentration. To avoid cross-contamination between samples as a consequence of evaporation/condensation and/or aerosols, reactions were prepared into 8-strip tubes with individually attached caps and covered with a drop of mineral oil (Sigma-Aldrich Company Ltd, UK). Amplifications were performed on Applied Biosystems® Veriti thermal cyclers (Life Technologies, CA, USA) with the following conditions: initial denaturation at 98°C for 5 min; 35 cycles of 98°C for 10 sec, 58°C for 20 sec and 72°C for 15 sec; final elongation at 72°C for 7 min. Samples and blanks as well as PCR negative controls (molecular grade water, *n* = 74) and positive

controls (genomic DNA [0.05 ng/µL] from *Maylandia zebra*, a cichlid from Lake Malawi not present in UK, *n* = 11) were amplified in triplicate. Amplicons were checked on 2% agarose gels stained with 10,000X GelRed Nucleic Acid Gel Stain (Cambridge Bioscience, UK). Gels were imaged using Image Lab Software (Bio-Rad Laboratories Ltd, UK) to visually check for contamination in blanks/PCR negative controls, presence of target band and consistency of results among replicates. After visualisation, PCR replicates were combined and samples belonging to the same collection date were pooled into sub-libraries using different volumes based on strength of PCR products on gels (no visible band = 20 µL, very faint or faint band = 15 µL, visible band =10 µL, bright band = 5 µL) (Alberdi *et al.* 2018). For each sub-library, 1 µL of the PCR positive controls and 10 µL of blanks/PCR negative controls were used. Sub-libraries were cleaned using a double-size selection magnetic beads protocol (Bronner *et al.*, 2009) with a ratio of 0.9X and 0.15X of magnetic beads (Mag-Bind® RXNPure Plus, Omega Bio-tek Inc, GA, USA) to PCR products (protocol details in Appendix 1). Two replicates of bead clean-up were performed per sub-library, and replicates were individually checked on a 2% agarose gel before pooling.

A second PCR, where Illumina tags were added to each sub-library, was then performed. Second PCRs were run in duplicate in a final reaction volume of 50 µL using 25 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 3 µL of each Illumina tag (10 µM; Integrated DNA Technologies, Belgium), 14 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 5 µL of cleaned sub-library. Second PCR thermal cycling conditions consisted of: initial denaturation at 95°C for 3 min; 8 cycles of 98°C for 20 sec and 72°C for 1 min; final elongation at 72°C for 5 min. Amplicons were checked on a 2% agarose gel where second PCR products were run alongside their non-tagged, cleaned counterparts to check for size differences after the addition of tags. A second double-size selection bead purification was carried out with a ratio of 0.7X and 0.15X of magnetic beads to PCR products (protocol details in Appendix 1). Tagged sub-libraries were quantified using the Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK) before being pooled at equimolar concentrations into a single final library. The final library was checked for size and integrity using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA), then quantified using the NEBNext® Library Quant Kit

for Illumina® (New England Biolabs® Inc., MA, USA). Following qPCR, a final dilution to 4nM was performed and 13 pM of the final denatured library was loaded onto the Illumina MiSeq® with 10% PhiX using 2 x 300 bp v3 chemistry (Illumina Inc., CA, USA).

Bioinformatics and statistical analyses

Raw sequence reads were demultiplexed using a custom Python script and then processed using metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.11 (https://github.com/HullUni-bioinformatics/metaBEAT), a bioinformatics pipeline incorporating commonly used open source software. Briefly, Trimmomatic v0.32 (Bolger *et al*., 2014) was used for read quality trimming (phred score Q30). During the trimming step, reads were also cropped to a maximum length of 110 bp and reads shorter than 90 bp were discarded. Additionally, to remove the locus primers, the first 18 bp of remaining reads were removed. FLASH v1.2.11 (Magoč and Salzberg, 2011) was then used to merge read pairs into single reads. For subsequent processing, merged reads and high quality forward only reads of sequences that failed to merge were kept. A final length filter (106 bp ± 20%) was applied to ensure sequences reflected the expected fragment size (106 bp). Remaining sequences were screened for detection of chimeric sequences against our custom reference database for UK fish (Hänfling *et al*. 2016) using the uchime algorithm (Edgar *et al*., 2011), as implemented in vsearch v1.1 (Rognes *et al*., 2016). Clustering at 100% identity in vsearch v1.1 (Rognes *et al*., 2016) was used to remove redundant sequences and possible sequencing errors as clusters represented by less than three sequences were omitted from further processing. Finally, retained reads were compared against a UK fish reference database (Hänfling *et al*. 2016) using BLAST (Zhang *et al*. 2000) and a lowest common ancestor (LCA) approach based on the top 10% BLAST matches for any query that matched a reference sequence across more than 95% of its length at minimum identity of 100%. Unassigned sequences from this comparison were subjected to a separate BLAST search against the complete NCBI nucleotide (nt) database using the same approach.

Final metaBEAT results were summarised as the number of reads assigned to each OTU in each sample screened, henceforth named as read counts. The final dataset with the complete fish OTU x read counts matrix was used to run downstream analyses

within R v3.5.1 (R Core Team, 2018). A low-frequency noise threshold of 0.001 (0.1%) was applied to the dataset to reduce the probability of false positives arising from cross-contamination or tag-jumping (De Barba *et al*. 2014; Hänfling *et al*. 2016). The threshold was applied using the proportion of reads assigned to each fish OTU calculated using the total read counts on a sample by sample basis. The choice of the threshold level was guided by the analysis of reads in positive controls that were not assigned to *M. zebra*. Furthermore, based on the low fish read counts found in sampling/filtration blanks and PCR negative controls, an additional threshold was applied, where all detections with less than 50 reads were removed.

Maps with circles proportional to Arctic charr read counts were used to visualise temporal and spatial patterns at the sites monitored. Shape files were read into R using the package *rgdal* (Bivand *et al*., 2019), and the *fortify* function together with the package *ggpolypath* (Sumner, 2016) were used to build the maps. All graphs were plotted using *ggplot2* (Wickham 2016).

3.4 Results

Sequencing outputs and bioinformatics

The sequencing run for the final library contained 266 samples in total, including 181 eDNA samples and 85 controls. The raw number of sequences generated was 41,160,110. An average of 73% sequences survived the quality trimming step, of which 98% were successfully merged. Following removal of chimera sequences and clustering, the total number of sequences for the library was 16,357,422. After taxonomic assignment against the 12S UK fish database (Hänfling *et al.*, 2016), 5,421,189 sequences matched 21 fish OTUs. All OTUs were identified to species level with the exception of sequences matching *Coregonus* and Percidae. The 12S region amplified cannot distinguish between two species of the family Percidae included in the database (*Perca fluviatilis* and *Sander lucioperca*) nor between two species belonging to the genus *Coregonus* (*C. lavaretus* and *C. autumnalis*; see also Hänfling *et al.*, 2016). Percidae records were manually assigned to *P. fluviatilis* as this is the only species represented in the study area.

Gill-netting survey

Arctic charr spawning individuals were caught and measured on eight out of 12 gill-netting surveys performed in autumn 2017. A total of 12 spawning Arctic charr were caught including: one ripe female, four spent or partially spent females, six running males and one spent male. Male specimens ranged from 25.3 to 31.5 cm, whereas female specimens ranged from 22.7 to 31.5 cm. On the 8th and 15th of November, nets were retrieved a few hours before the eDNA sampling, and one ripe female (30.6 cm) and one spent female (22.7 cm) were caught respectively. In addition to Arctic charr spawning individuals, three more fish species (*R. rutilus*, roach; *Esox lucius*, pike; *P. fluviatilis*, European perch) were caught at SNTH during the twelve netting surveys.

Variation in Arctic charr eDNA signal

Overall 155,678 reads were assigned to Arctic charr. In the pre-spawning period (one sampling event in October 2017), we did not find any reads assigned to Arctic charr at neither SRN nor SNTH (Fig. 2). During the spawning period (between November and December 2017), reads were assigned to Arctic charr on all eight sampling events at both locations (Fig. 2, Appendix 2: Fig. S1). The number of reads assigned to Arctic charr in the spawning period ranged from 30,735 to 188 at SRN (shoreline, putative spawning grounds) and from 15,263 to 419 at SNTH shore (demonstrated spawning grounds). Arctic charr reads were also detected in the deepest waters along the offshore transect (15,695 to 296 at OF1 to OF5; Fig. 2, Appendix 2: Fig. S1).

The highest read counts for Arctic charr were found at SRN on the shoreline transect in December 2017 (30,735 reads at site SRN4; Fig. 2), and 23 of 55 samples collected at SRN during the spawning period were positive for Arctic charr. The highest occupancy and read counts were found in December (7/8 sites, with the highest read counts 30,735 and 13,949 reads at site SRN4 and SRN8 respectively, Fig. 2). At the demonstrated spawning grounds of SNTH shore (SNTH1, SNTH2, SNTH3; Fig. 1), the highest read counts of Arctic charr (15,263 reads) were observed at site SNTH2 on 13th November (Fig. 2). Overall, all samples collected in autumn showed positive detection of Arctic charr with the exception of sampling site SNTH1 on 17th November (Fig. 2). At the deepest sites along the offshore transect (site OF1 to OF5; Fig. 1) the highest read

counts matching Arctic charr were found on 8th November at site OF3 (15,695 reads; Appendix 2: Fig. S1). Across offshore sites, 18 samples out of 25 collected showed positive Arctic charr detection across the spawning period (Fig. 2, Appendix 2: Fig. S1).

In January (after spawning), we observed a reduction of the eDNA signal assigned to Arctic charr at both transects (Fig. 2). Only two sites at the lake side (SRN5 and SRN6) showed positive detection of Arctic charr with low read counts (584 and 489 reads) and, two more offshore sites (OF2, OF5) with higher number of reads (916 and 1003 reads; Fig. 2). In spring, water collection occurred in April and only one sample belonging to the offshore transect (OF3) showed positive detection of Arctic charr with 1844 reads assigned (Fig. 2). The reduction in Arctic charr eDNA signal culminated in July during our last eDNA sampling campaign, when no detections of Arctic charr were found at any sites sampled (Fig. 2).
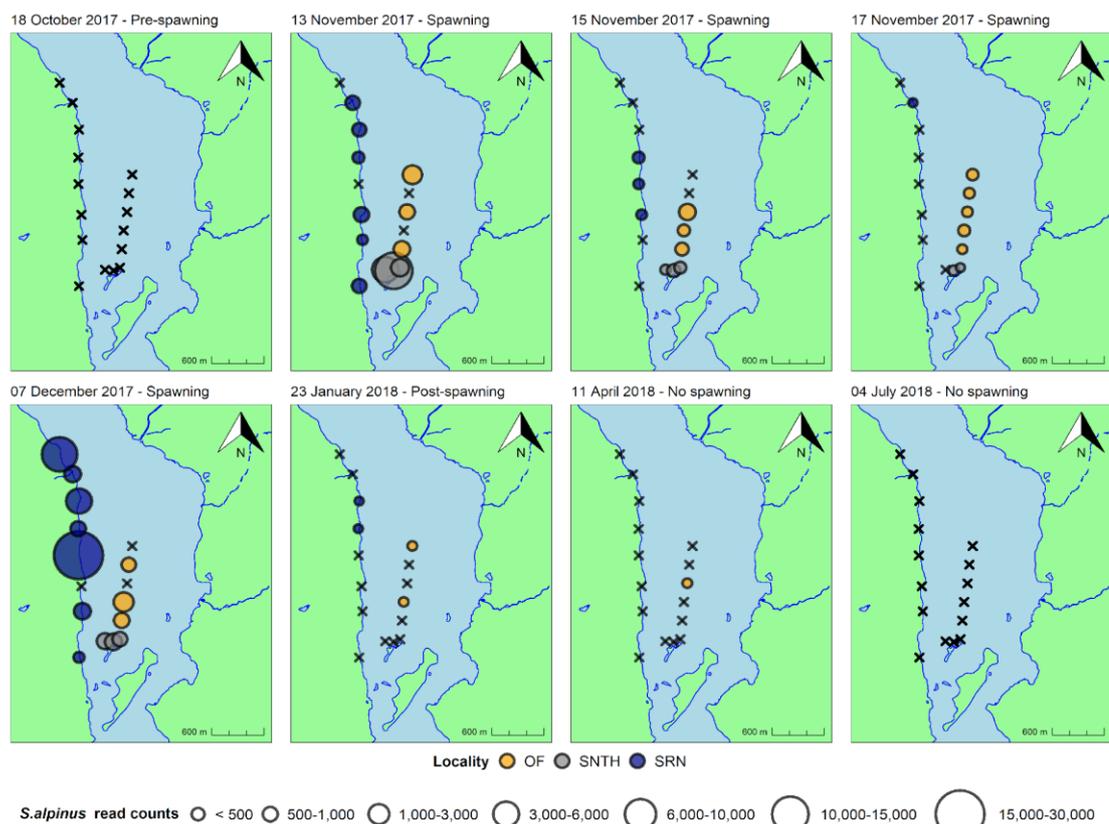
Figure 2. Spatio-temporal variation of Arctic charr eDNA signal in the north basin of Lake Windermere. Bubble size is proportional to read counts assigned to Arctic charr, whereas black crosses indicate sites where the species was not detected. Only eDNA sampling events where collection occurred at all localities are represented in the maps (see Appendix 2: Fig. S1 for results of additional sampling events). Maps were created using shape files downloaded from EDINA Digimap® Ordinance Survey service (http://edina.ac.uk/digimap).

Fish species diversity and community analyses

A total of 18 OTUs were identified across all sampling locations and dates. The six most common fish species detected consistently with high read counts across all sampling dates and localities were: *Anguilla anguilla* (European eel), *Cottus gobio* (European bullhead), *E. lucius* (pike), *P. fluviatilis* (European perch), *R. rutilus* (roach), *Salmo trutta* (brown trout) (Fig. 3). Three species showed a lower but still high number of detections and read counts across different dates: *Abramis brama* (common bream), *Phoxinus phoxinus* (Eurasian minnow) and Arctic charr (Fig. 3). Five fish species were considered rare given the low number of detections and reads assigned: *Barbatula barbatula* (stone loach), *Gasterosteus aculeatus* (three-spined stickleback), *Lampetra fluviatilis* (European river lamprey), *Gymnocephalus cernua* (ruffe), *Salmo salar* (Atlantic salmon) (Fig. 3; Appendix 2: Fig. S2). Lastly, four taxa showed a maximum number of three detections across all sites and dates: *Platichthys flesus* (flounder), *Oncorhynchus mykiss* (rainbow trout), *Osmerus eperlanus* (European smelt) and *Coregonus* spp*.* (coregonids)*.*

Species such as eel, stone loach, bullhead, European river lamprey, or Eurasian minnow were more often observed in shoreline sites (SRN, SNTH; Fig. 3), while common bream, Atlantic salmon, pike typically showed higher detection in offshore sites (OF; Fig. 3). Overall, the number of fish species detected per locality at different dates varied from a maximum of 13 (13th November, SRN) to a minimum of five (4th July, OF) species. The lowest species richness was found during the months of October and July, where the number of fish species detected per locality ranged between five and nine (Fig. 3). Between the eDNA sampling campaigns on 15th November and 23rd January, little variation in fish species richness was found across different localities (Fig. 3).
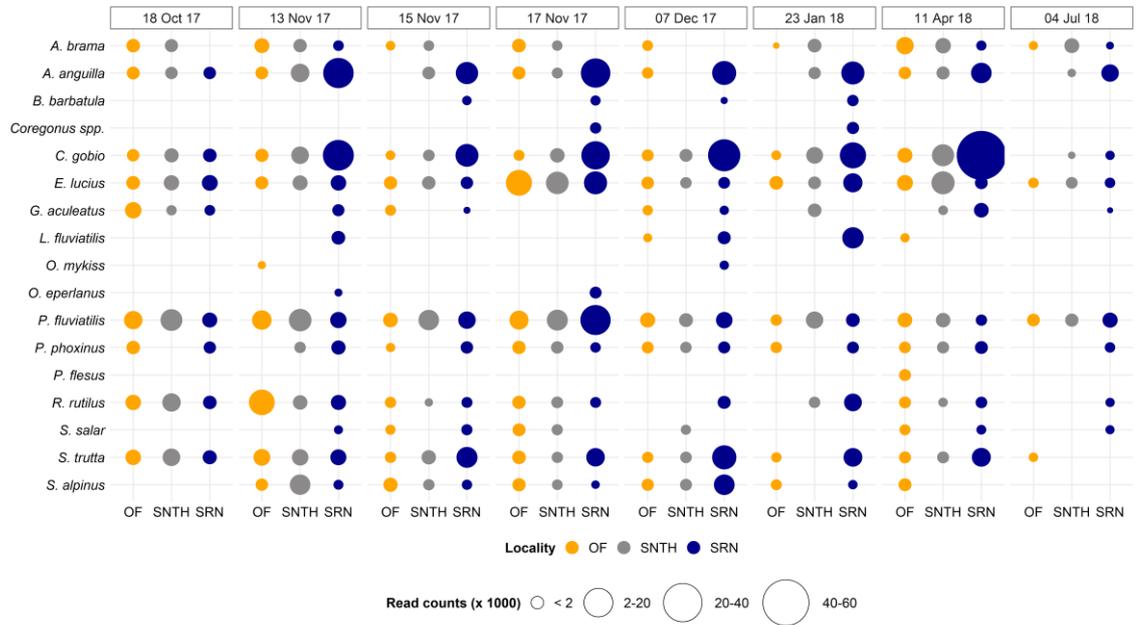
Figure 3. Fish community plot of eDNA metabarcoding detections and relative abundance. Bubble plot showing average read counts of detected fish species across different eDNA sampling dates and localities in Windermere's north basin. SRN includes average read counts of species detected at the shoreline sites of Red Nab (n = 8), SNTH includes sites at the shore of North Thompson island (n = 3) and OF includes offshore sites between 20 and 40 m depth (n = 5). See Appendix 2 (Fig. S2) for fish community results of the additional sampling events.

3.5 Discussion

The present study revealed the ability of eDNA metabarcoding to monitor the reproductive activity of Arctic charr autumn-spawning individuals in Windermere. Supporting the evidence from our molecular observations were catches of mature specimens at the monitored breeding sites where peaks of Arctic charr read counts were detected. In addition, the temporal gradient of the genetic signal observed in autumn was a further indication of the species spawning activity. In this study, we have characterised times and locations of Arctic charr spawning events, revealing key information on putative breeding localities where spawning has not been monitored or observed for over 50 years.

Arctic charr eDNA is absent at the monitored sites outside the spawning season

In large lentic systems, the organisms' genetic signal is localised, and the spatial and temporal distribution of eDNA resembles the sites occupied by a species in the water

at a given time (Zhang *et al.*, 2020; Lawson Handley *et al.*, 2019). Previous fish eDNA surveys in Windermere showed that Arctic charr is mainly detected in deep waters outside the species' spawning season (Lawson Handley *et al.*, 2019; Hänfling *et al.*, 2016). In agreement with these studies, the species biology, and our initial hypotheses, we observed seasonally-limited detections of Arctic charr at the localities monitored. Arctic charr eDNA was not detected in water samples collected pre-spawning (October; Fig. 2) and post-spawning (April-July; Fig. 2) at the shoreline locations of SRN and SNTH (putative and monitored breeding grounds respectively).

Arctic charr feeding grounds are located in the offshore areas of Windermere (Frost, 1977; Mills 1990); however, the sites located along the depth transect (OF; Fig. 1), and sampled outside the breeding season, showed no detection of Arctic charr (October, April and July; Fig. 2). The limited sampling effort carried out in the deep waters of the lake may have hindered the detection of Arctic charr at the species' feeding grounds. A more comprehensive sampling effort along the lake midline would have found the species in the deepest areas of the lake beyond its spawning season as shown in Lawson Handley *et al.*, 2019. The seasonal stratification occurring during the warmer months limited the dispersion of Arctic charr DNA, which remained confined to the deepest feeding areas, and explain the absence of Arctic charr detections at the sites surveyed during the warm season (Lawson Handley *et al.*, 2019).

Arctic charr eDNA is detected consistently at the monitored sites in autumn

Consistent detections of Arctic charr at all localities surveyed (OF, SRN and SNTH; Fig. 2) were found during the species' spawning season between November and December. Previous assessments of Arctic charr autumn-spawning population in Windermere described the species' breeding grounds at shallow depths (less than 5 m) (Frost, 1965; Miller *et al.*, 2015; Winfield *et al.*, 2015). Detections of Arctic charr eDNA at deep water sites during the breeding season might be explained by migratory mature individuals moving from the offshore feeding grounds to the shallow breeding habitats in autumn. Migratory patterns to the spawning grounds in fish have been inferred before using eDNA concentrations (Thalinger *et al.*, 2019; Antognazza *et al.*, 2019; Erickson *et al.*, 2016).

During the spawning season, the highest number of reads for Arctic charr were found at the shallow breeding habitats of SRN and SNTH (Fig. 2). In agreement with previous assessments of the suitability of these autumn-spawning grounds to support Arctic charr reproductive activity (i.e. catches; visual surveys, hydroacoustic, PAM; Miller *et al.*, 2015; Winfield *et al.*, 2015; Bolgan *et al.* 2017), our results found other evidence of the species' spawning activity at these shallow localities of the lake. At SNTH, catches of mature Arctic charr individuals indicative of spawning activity during the netting survey were associated to peaks in read counts found with our eDNA survey. As the same or higher Arctic charr eDNA signal was detected at the putative spawning grounds of SRN only during the spawning season, we infer that spawning activity was occurring at this locality even though mature individuals have not been caught at these sites in the last 50 years. Ethical implications of destructive established methods (i.e. netting) restrict the application of these monitoring approaches, especially when the target species are threatened, such as Arctic charr in Windermere (Winfield *et al.*, 2009; Winfield, 2002). Therefore, we have demonstrated the suitability of eDNA metabarcoding as a universal, non-invasive molecular tool to infer spawning activity through the temporal and spatial localisation of the Arctic charr genetic signal in a large lake.

Alongside the localised genetic signal of Arctic charr within the breeding season, eDNA peaks in the species' read counts were observed in autumn at two sites of SNTH and SRN (15,263 reads at OF2 and 30,735 SRN4; Fig. 2). While the time-limited eDNA signal of Arctic charr at the shoreline breeding grounds (SRN and SNTH) can be associated with several reproductive features of the species at these breeding areas (i.e. nesting females, courting and/or competing males, aggregation of mature individuals), the peaks of eDNA observed here are most likely associated with the release of gametes. Laboratory and field eDNA studies have shown the association between peaks in eDNA concentrations and gamete release by aquatic organisms (Takeuchi *et al.*, 2019; Bylemans *et al.*, 2017)

Fish communities at the monitored sites and interaction with Arctic charr

During the twelve netting surveys carried out at SNTH between November and December within the Arctic charr monitoring programme, three additional fish species were caught at those shallow breeding sites: *R. rutilus*, *E. lucius*, and *P. fluviatilis*. Our five eDNA metabarcoding surveys performed within the same period at SNTH (3 sites; Fig. 1) detected six additional species including *S. salar*, *S. trutta*, *P. phoxinus*, *C. gobio*, *A. anguilla* and *A. brama* (Fig. 3; Appendix 2: Fig. S2). The higher number of species detected using eDNA metabarcoding as opposed to established methods is in agreement with the increasing number of studies that have observed higher detection rates of fish species with eDNA metabarcoding compared to established methods (Valentini *et al.*, 2016; Civade *et al.*, 2016; Hänfling *et al.*, 2016).

Of the 18 fish OTUs detected within our temporal eDNA metabarcoding survey in Windermere's north basin, 14 species are known to occur and have been consistently detected in previous eDNA surveys of the lake (Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019). Four additional taxa were found with the lowest number of detections (*O. eperlanus*, *O. mykiss*, *P. flesus* and *Coregonus* spp.). When eDNA records are not supported by historical evidence of species occurrence, inconsistent detections of organisms' DNA in the environment can reflect true records of low abundant taxa or false positive detections of alternative sources of DNA (Jerde, 2019). *Osmerus eperlanus*, *O. mykiss*, *Coregonus* spp. and *P. flesus* were never recorded in Windermere, but these taxa are known to occur in the lake's catchment and, eDNA transport from other areas of the catchment could explain the DNA detections in the lake (Hänfling *et al.*, 2016). The former three species are also commonly used as dead-bait from pike anglers, particularly active in Windermere in the autumn-winter seasons, consistent with the limited time frame in which eDNA detections of these species were found (Fig. 3; Appendix 2: Fig. S2). According to the authors' knowledge, hybrids of such taxa do not occur in Lake Windermere, however the hypothesis of hybridisation, which could explain the origins of the taxa mitochondrial DNA in the lake, cannot be excluded, and it has been documented in other freshwater habitats especially between salmonids and other fish (e.g. *O. mykiss;* Muhlfeld *et al.*, 2014). The presence of low-abundant populations of *O. mykiss, O. eperlanus, Coregonus* spp. and

P. flesus within the lake should also not be ruled out as former eDNA surveys in Windermere demonstrated the ability of this tool to early-detect unrecorded fish species, such as the case of *G. cernua*. In this study a single detection of *G. cernua* was found with relatively high read counts at one offshore site in November (7473 reads; Appendix 2: Fig. S2). This species was detected for the first time in Windermere by Hänfling *et al.*, 2016 during a more comprehensive eDNA survey of the lake. Given that *G. cernua* has not been recorded from conventional surveys in Windermere until 2019, researchers could not exclude eDNA transport from upstream populations or other sources of the species DNA (Hänfling *et al.*, 2016). Recent findings of the Centre for Ecology and Hydrology's long-term monitoring efforts for fish communities in Windermere (September 2019) confirmed the presence of *G. cernua* in the lake, which in combination with a number of other studies, demonstrates the power of eDNA metabarcoding for early detection of non-native species (Blackman *et al.*, 2017; Holman *et al.*, 2019).

The recent introduction of *G. cernua* in Windermere is likely to cause increasing concern given the species' potential for causing changes in fish species composition in inshore and offshore habitats as a result of predation or competition as observed in other UK lakes (Winfield *et al.*, 2011). Windermere has experienced species introductions in the past, for instance, *R. rutilus* and *A. brama* are now established in the lake and their expansion has been accompanied by a decrease in Arctic charr abundance, especially in the south basin (Winfield *et al.*, 2011; Winfield *et al.*, 2008a). As it is clear that Windermere is undergoing changes affecting the native fish community, including Arctic charr, continued vigilance of researchers, government agencies, and the public regarding introduced species and the assessment of native populations will become paramount, and sensitive monitoring approaches such as eDNA metabarcoding can support these conservation efforts.

Our eDNA metabarcoding results additionally detected *L. fluviatilis*, one of the rarest species in the lake, which previous eDNA surveys likely did not detect because of the restricted and localised spatial and temporal distribution of this species in Windermere (see Fig. 3; Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019). As opposed to these rare detections, the genetic signal of *C. gobio* appeared very robust and consistent at

the monitored sites. The territorial behaviour of *C. gobio,* and the preference of gravel or stony bottoms for hiding strategies against predation, are likely the reasons for consistency in its genetic footprints at the grounds surveyed. In addition, the eDNA peak observed in April could be associated with the species' spawning activity, occurring between March and April (TomLinson and Perrow, 2003; Smyly, 1957). As opportunistic feeders, *C. gobio* individuals are also known to prey on salmonid eggs, another possible justification for their high detections at Arctic charr spawning grounds, especially the ones located on the west shore of the lake (SRN; Gaudin and Caillere, 2000; Palm *et al*., 2009; Smyly, 1957). Another species with a strong genetic signal at the sites monitored was *A. anguilla*. *Anguilla anguilla* is one of the most widely distributed species in Windermere and previous eDNA surveys have shown its uniform detection within the lake across different seasons and transects (Hänfling *et al*., 2016; Lawson Handley *et al*., 2019). The high *A. anguilla* eDNA signal at the shore of Red Nab could be also associated with the species' predation behaviour on Arctic charr eggs as indicated by Frost (1952) who found Arctic charr eggs in *A. anguilla* stomachs. This represents another possible interaction with the focal species of this study.

3.6 Conclusion

Our eDNA fish community observations at Arctic charr spawning grounds in Windermere's north basin unveiled the ability of eDNA metabarcoding to monitor Arctic charr reproductive activity. Furthermore, we have demonstrated, in line with other studies, that this non-invasive whole-community approach outperforms established methods for fish community survey in freshwater environments (Hänfling *et al*., 2016; Valentini *et al*., 2016; Civade *et al*., 2016). In agreement with our previous eDNA observations (Hänfling *et al*., 2016), we additionally confirmed that this tool has the capacity of providing information on critical detection of non-native species at the early stages of colonisation. As suggested by Jerde (2019), monitoring programmes should have greater confidence in eDNA findings and overcome the uncertainty of "not having a fish in hand" whenever a consistent genetic signal is observed. This will allow policy makers and monitoring agencies to take prompt action for the implementation of suitable conservation and eradication efforts. This research entails novel

applications of eDNA metabarcoding, enhancing the tool's capabilities far beyond its qualitative and semi-quantitative capacities. Our observations showed that this non-invasive approach provides reliable information on species' reproductive events, thereby contributing to the assessment of fish populations through recruitment monitoring, an essential aspect for conservation and population management.

Chapter 4

Targeted or whole-community? A "sensitive" matter for environmental DNA samples.

Arctic charr from Lake Windermere (Cumbria, UK). Photo credit: Dr. Winfield I.J., used with permission.

4.1 Abstract

The sampling of environmental DNA (eDNA) coupled with species-specific or whole-community molecular tools has the ability to overcome some of the issues associated with current fish monitoring surveys. A number of studies have compared the sensitivity of species-specific and community-wide methods for the detection of low-abundant or invasive species in eDNA samples. Such studies have concluded that species-specific molecular approaches should be considered when one species is the target as they are more sensitive than community-wide methods, however, these comparisons often have not used an adequate experimental design (e.g. different number of PCR replicates between methods). In the present study we aim to compare the sensitivity of molecular approaches for the detection of Arctic charr (*Salvelinus alpinus* L.) in water samples using a species-specific qPCR assay and a metabarcoding protocol with vertebrate-specific primers. Twenty-three eDNA samples were analysed using both molecular methods and 12 PCR replicates. Overall, we observed a similar number of positive detections between methods, and across all samples and replicates

(66% for metabarcoding and 61% for qPCR), and found a positive and significant correlation between the methods detection probability (r= 0.66, p < 0.000). Furthermore, we used an occupancy model to investigate the influence of original eDNA concentration and 12S eDNA concentration (region used for metabarcoding analysis) on Arctic charr detection in PCR replicates and found that both factors play a significant and positive role. This study highlights that metabarcoding can be as sensitive as qPCR when a similar technical replication effort is used, hence, supporting the use of wide-community approaches for community analyses as well as for sensitive detections of low-abundant species in complex environmental mixtures.

4.2 Introduction

Accurate species detection together with distribution and abundance estimates are the key objectives of biomonitoring surveys. Data of species distribution are therefore crucial to improve our understanding of ecological communities and networks, and to provide information to environmental managers and policy-makers for the evaluation of conservation and management actions of low-abundant and endangered species (Stork and Samways, 1995). However, imperfect detections of rare taxa that underestimate species richness and distribution may lead to poor management decisions (Schmidt *et al.*, 2013; MacKenzie *et al.*, 2002). Rate of imperfect detections can be higher for aquatic taxa and the efficiency of monitoring methods depend on the target species, their size, abundance and developmental stage (Lintermans, 2016). For example, established fish survey methods (i.e. visual surveys, netting/trapping) have a low detection probability of small, elusive and low-abundant species and can be locally limited by taxonomic identification of cryptic species (i.e. in sub-tropical and tropical regions; Port *et al*., 2016; Nester *et al.*, 2020).

The sampling of environmental DNA (eDNA) from aquatic environments coupled with species-specific or whole-community molecular tools has the ability to overcome some of the issues associated with current fish monitoring surveys. The non-invasive collection of water samples can be extensively applied across a variety of aquatic habitats without specific logistic or ethic restrictions allowing higher species' detection probability (see Ruppert *et al.*, 2019 for an overview). Furthermore, molecular analyses

of the genetic material in environmental samples allows the reliable identification of specific taxonomic units when reference sequences in public databases are available (Hänfling *et al.*, 2016; Hering *et al.*, 2018; Pont *et al.*, 2019). Sensitivity, in the context of eDNA surveys, is the probability of detecting target DNA at a site where that DNA is present. Estimating sensitivity, or probability of detection, is important for interpreting the results of eDNA-based studies because it specifies the chance of detection failure, or committing a false negative or type II error (Furlan *et al.*, 2016). A number of biological, environmental, technical and methodological factors influence the molecular detection of a taxon in environmental samples. For instance, species biology, life stage and ecology together with population level dynamics such as abundance, density and migratory patterns determine a variation of detection probability of target species (de Souza *et al.*, 2016; Erickson *et al.*, 2016; Stoeckle *et al.*, 2017b). Environmental variables (e.g. UV radiation, temperature, pH) and physical changes, such as horizontal and vertical transport in water, additionally contribute to the variation of detection probability (see Barnes *et al.*, 2016 for an overview). Detection probability in eDNA surveys also depends on the level of in-field replications (number of samples collected) or in-laboratory replications (filtration, extraction or PCR replicates) or other factors such as inhibition, sequencing depth for community assays or the sensitivity of the molecular assay used (Chapter 2; Piggot, 2016; Schmidt *et al.*, 2013; Smith and Peay, 2014).

High sensitivity is perceived to be a key advantage of eDNA methods with a number of comparative studies showing higher sensitivity of species-specific assays compared to whole-community approaches (Bylemans *et al.*, 2019; Harper *et al.*, 2018; Lacoursière-Roussel *et al.*, 2016c; Wood *et al.*, 2019). Targeted approaches are thought to be more sensitive as they are developed and optimised around the specific detection of target DNA (Lodge *et al.*, 2012; Laramie *et al.*, 2015; Dougherty *et al.*, 2016; Yusishen *et al.*, 2020), whereas the community approach, namely metabarcoding, aims to characterise entire species assemblages, hence sacrificing specificity in favour of taxonomic diversity (Bylemans *et al.*, 2018; Wood et al., 2019; Marques *et al.*, 2020). The use of broad-range primers in metabarcoding analyses may lead to amplification biases such as the preferential amplification of DNA from very abundant species or DNA sequences with lower primer-template mismatches within the mixture (Elbrecht and Leese, 2015;

Bylemans *et al.*, 2018). Amplification biases can ultimately increase the chance of false-negative detections compared to a targeted approach. However, comparative eDNA studies between targeted and metabarcoding approaches are few and have not used the same number of PCR replicates or have used primers amplifying regions of different lengths (Harper *et al.*, 2018; Lacoursière-Roussel *et al.*, 2016c). If community-wide and targeted eDNA approaches are shown to have comparable sensitivity in the detection of certain species in specific habitats, the use of metabarcoding could be advantageous because the community information can play a key role in the evaluation of the current status of certain population or even predict future changes in community structure due, for example, to the early identification of biological invasions (Blackman *et al.*, 2017; Borrel *et al.*, 2017; Klymus *et al.*, 2017).

Cost and time efficiency is also an essential aspect that should be considered when planning a monitoring or conservation programme. Harper *et al.*, 2018 extensively compared cost and investigation effort of qPCR and metabarcoding and found that whilst the investigator effort between molecular methods was comparable, metabarcoding was slightly more expensive than qPCR. It should be noted however that, in Harper *et al.*, 2018, the technical replication effort was higher in qPCR (12 PCR replicates) than in metabarcoding (three pooled PCR replicates), hence metabarcoding would be much more expensive if the same level of technical replication between methods is achieved.

In this study, we evaluated the sensitivity of targeted qPCR and metabarcoding for the detection of Arctic charr (*Salvelinus alpinus* L.) in eDNA samples from Lake Windermere. Here, we used the same number of PCR replicates (*n*=12) for both experiments and maximised the sequencing depth by sequencing each PCR replicate separately in order to achieve adequate fairness in the comparison of the molecular assays used. We additionally estimated the original total DNA concentration and 12S DNA concentration (region amplified in metabarcoding analysis) within our samples to investigate the influence of these factors in detection probability of Arctic charr in PCR replicates.

4.3 Materials and methods

Sample collection and processing

A subset of extracted eDNA samples collected at Windermere (Lake District, Cumbria, UK) during the Arctic charr spawning season and analysed in Chapter 3 was used for this study. Sample selection was informed by previous metabarcoding analyses and sample extracts were chosen to represent a wider range of read counts assigned to Arctic charr (ranging from 0 to over 30,000 reads; Fig. 1). A total of 23 extracts, collected at 16 different locations and representing three different habitats, were used including: eight shoreline samples (SRN in Chapter 3) collected on the 7th December 2017 from the west shore of the lake North Basin; seven offshore samples collected on the 8th November 2017 and eight offshore samples collected on the 13th November 2017 (Fig. 1). Detailed information on sampling, filtering and extraction protocols can be found in Chapter 3.
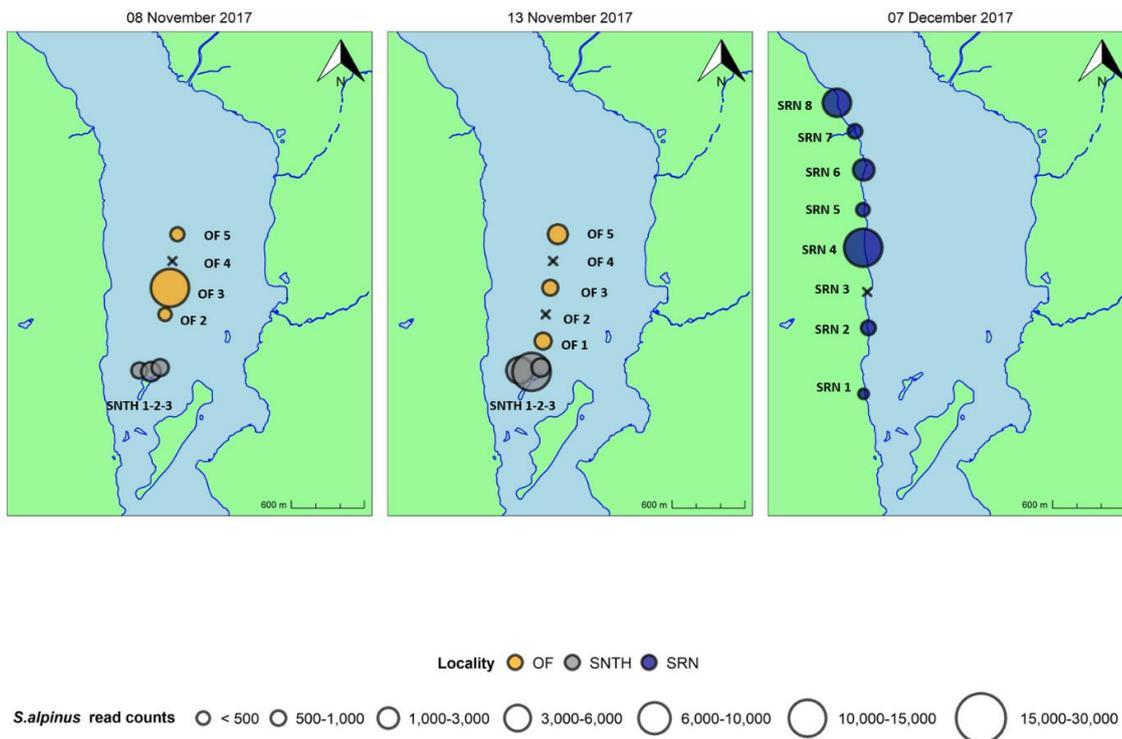

Figure 1. Bubble map of Windermere eDNA samples chosen for this study. The 23 eDNA samples selected for the analyses presented in this study were chosen based on previous metabarcoding analysis (Chapter 3) and represented a wide range of read counts (ranging from 0 to over 30,000 reads) assigned to Arctic charr as shown in the figure.

Metabarcoding and qPCR experiments were carried out within the same time frame, and prior to perform these molecular analyses, the total nucleic acids yield and purity of the 23 eDNA extracts was checked using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) and the double-stranded genetic material was quantified using the Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK).

12S DNA quantification

To determine the influence of 12S DNA concentration on Arctic charr detections in PCR replicates this fragment was quantified using a SYBRGreen qPCR assay with 12S primers targeting a 106 bp region (Riaz *et al.* 2011; Kelly *et al.* 2014), the same region used for metabarcoding. Quantitative PCR reactions were prepared using 10 μL of iTaq™ Universal SYBR® Green supermix 2X (Bio-Rad Laboratories, CA, USA), 0.40 μL of each primer (10 μM; Integrated DNA Technologies, Belgium), 7.20 μL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 μL of DNA template. Reactions were run on 96-well plates using StepOnePlus™ Real-Time PCR system (Life Technologies, CA, USA). Cycling conditions included: initial denaturation of 3 min at 95 °C and 40 cycles of 30 sec at 95 °C and 30 sec at 57 °C; melting curve step from 65 °C to 95 °C with a temperature increment of 0.5 °C for 10 sec. Samples were run in triplicates and each plate included three negative controls (i.e. no DNA template added). For the relative quantification of 12S DNA, the standard curve was built using five 10-fold dilutions with five replicates of a fish mock community with the genomic DNA of five fish species (pike, *Esox lucius*; Eurasian minnow, *Phoxinus phoxinus*; pumpkinseed, *Lepomis gibbosus*; topmouth gudgeon, *Pseudorasbora parva*; rudd, *Scardinius erythrophthalmus*) diluted to 5 ng/μL. Dilutions were checked using the Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit and then combined and diluted again to a final concentration of 1 ng/μL representing the first point of the standard curve.

Arctic charr qPCR assay

Arctic charr primers and probe were designed visually in AliView v.1.20 (Larsson, 2014). Species-specific primers (forward primer: 5'-GATCCTCCACACCTCTAAACAG-3'; reverse primer: 5'-GGCAACTTGGCCGATAATGATA-3') and a locked nucleic acid probe (5'-56-FAM-AACCCT+GGTAGC+AGACATACTA-3BHQ_1-3') targeting a 148 bp region of

the mitochondrial DNA cytochrome B (cytB) gene were selected. Primers and probe parameters (i.e. melting temperatures, hairpins, self-dimers and hereto-dimers) were checked using the online software OligoAnalyzer Tool (https://www.idtdna.com/pages/tools/oligoanalyzer) and Beacon Designer (http://www.premierbiosoft.com/qOligo/Oligo.jsp?PID=1). Primers specificity (without probe) was tested *in silico* against the National Centre of Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/) nucleotide database with BLASTn (Basic Local Alignment Search Tool) and ecoPCR (Ficetola *et al.*, 2010) against a custom-curated reference database of 67 UK freshwater fish species (Hänfling *et al.*, 2016).

In vitro testing of the assay was performed using qPCR. Amplification results were checked using genomic DNA extracts (standardised to 1 ng/μL) of Arctic charr and four closely related species occurring in UK lacustrine environments (brown trout, *Salmo trutta*; Atlantic salmon, *Salmo salar*; rainbow trout, *Onchorhynchus mykiss*; brook trout, *Salvelinus fontinalis*). The assay was also tested on field eDNA samples from UK lakes where the target species is known to be present or absent (i.e. Middle Lake, Chapter 2; Grimwith Reservoir and Loch Earn, Chapter 5). The qPCR assay was optimised using tissue extracted from Arctic charr on a gradient of annealing temperatures. A licence to use a gillnet to capture Arctic charr in Loch Earn for the purposes of this research was provided by Marine Scotland Science (Licence CSM-16-136). The limits of detection (the concentration at which no Arctic charr DNA amplifies) and quantification (the concentration at which Arctic charr DNA amplifies inconsistently across replicates) were determined using 10-fold dilutions (5×10^6 to 5 copies/μL) of a 400 bp gBlocks® Gene Fragment (Integrated DNA Technologies, Belgium) based on GenBank accession (AF154851.1) for Arctic charr. Copy number for the gBlocks® fragment was calculated following the manufacturer's instructions.

Before qPCR analysis, the 23 eDNA samples were tested for the presence of PCR inhibitors using the "C_t shift" method (Jane *et al.*, 2015; Tillotson *et al.*, 2018) and two types of controls: an exogenous internal positive control (EXO-IPC) performed in duplicate on each DNA extract (Applied Biosystems™ TaqMan™ Exogenous Internal Positive Control assay; Fisher Scientific, UK) and an internal positive control consisting in 2 μL of Arctic charr genomic DNA (1 ng/μL) spiked into the reaction and performed

in triplicate. The C_t value defines the number of cycles required to the PCR product to cross the threshold recognised by the qPCR instrument, and PCR inhibitors in eDNA samples will shift the C_t during the amplification. Here, we considered a shift of the mean C_t values (across replicates) greater than two compared to the C_t values of control samples (no eDNA template) as indication of inhibition. Specifically, IPC assay consisted of 10 μL TaqMan® Environmental Master Mix 2.0 (Life Technologies, CA, USA), 2 μL EXO-IPC mix and 0.4 μL EXO-IPC DNA, 5.6 μL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 μL eDNA template per 20 μL total volume reaction. Cycling parameters consisted of initial steps of 2 min at 50 °C then 10 min at 95 °C, followed by 40 cycles of denaturing at 95 °C for 15 sec and annealing/extension at 60 °C for 1 min. Inhibition tests with spiked Arctic charr genomic DNA (1 ng/μL) were performed using the Arctic charr assay developed in this study, and described below, reducing the volume of molecular grade water to compensate the 2 μL of genomic DNA added to the reactions.

For our qPCR analysis, 12 PCR replicates were performed using 10 μL of TaqMan® Environmental Master Mix 2.0 (Life Technologies, CA, USA), 1 μL of each primer (450 nM; Integrated DNA Technologies, Belgium), 1 μL of LNA-probe (125 nM; Integrated DNA Technologies, Belgium), 5 μL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 μL of DNA template (same as in the sequencing library preparation). Reactions were run on 96-well plates using StepOnePlus™ Real-Time PCR system (Life Technologies, CA, USA). Cycling conditions included: initial incubation of 2 min at 50 °C; initial denaturation of 10 min at 95 °C and 45 cycles of 15 sec at 95 °C and 60 sec at 58 °C. A total of five qPCR plates were run to accommodate a maximum number of five samples with 12 PCR replicates in each plate, three negative controls (i.e. no DNA template added) and a standard curve with five dilutions and five replicates for each standard. DNA extracted from Arctic charr tissue was quantified with Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK) and used to generate the standard curve using five 10:1 serial dilutions. Concentrations for the serial dilution ranged from 10 ng/μL to 1×10^{-3} ng/μL and five replicates of each dilution were used per plate.

All qPCR replicates of the 23 eDNA extracts were used for downstream analyses including the technical replicates with C_t differences > 0.5 C_t, the commonly accepted range of variation for accurate quantification.

Metabarcoding

Library preparation included a two PCR and a nested-tagging approach as described in Li *et al.* (2019a). Briefly, 12 rounds of PCR 1 (12 PCR replicates) were used to amplify each of the 23 eDNA samples using unique indexed primers targeting a 106 bp region within the 12S fragment (Riaz *et al.* 2011; Kelly *et al.* 2014). At each round of PCR, a negative (molecular grade water) and a positive control (genomic DNA from *Maylandia zebra*, a cichlid not occurring in the UK) were included. PCRs were performed within a final reaction volume of 25 µL including 12.5 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 1.5 µL of each indexed primer (10 µM; Integrated DNA Technologies, Belgium), 0.5 µL of the Thermo Scientific™ Bovine Serum Albumin (Fisher Scientific UK Ltd, UK), 7 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 2 µL of DNA template at the original sample concentration. Amplifications were performed on Applied Biosystems® Veriti thermal cyclers (Life Technologies, CA, USA) with the following conditions: initial denaturation at 98°C for 5 min; 35 cycles of 98°C for 10 sec, 58°C for 20 sec and 72°C for 15 sec; final elongation step at 72°C for 7 min. Amplicons were checked on 2% agarose gels stained with 10,000X GelRed Nucleic Acid Gel Stain (Cambridge Bioscience, UK). After visualisation, samples belonging to the same amplification round of PCR1 (23 eDNA samples and control samples) were pooled and normalised using different volumes as deduced from strength of PCR products on gel (very bright = 5 µL, bright = 10 µL, faint = 15 µL, and very faint/no band = 20 µL) using 1 µL of the positive control and 10 µL of the negative control for each pool (Alberdi *et al.* 2018). Pooling volumes were consistent across different sub-libraries and 12 sub-libraries, corresponding to the 12 PCR 1 replicates, were prepared.

Sub-libraries were cleaned using a double-size selection magnetic beads protocol (Bronner *et al.*, 2013) with a ratio of 0.9X and 0.15X of magnetic beads (Mag-Bind® RXNPure Plus, Omega Bio-tek Inc, GA, USA) to PCR products. Illumina tags were added to each cleaned sub-library with a second amplification round. Second PCRs were run

in duplicate in a final reaction volume of 50 µL using 25 µL of Q5® Hot-Start High-Fidelity 2X Master Mix (New England Biolabs® Inc., MA, USA), 3 µL of each Illumina tag (10 µM; Integrated DNA Technologies, Belgium), 15 µL of molecular grade water (Fisher Scientific UK Ltd, UK) and 4 µL of templates. Second round PCRs consisted of: 95°C for 3 min; 8 cycles of 98°C for 20 sec and 72°C for 1 min; and 72°C for 5 min. PCR products were checked on a 2% agarose gel alongside their non-tagged products to check for size differences after tag addition. A second double-size selection bead purification was carried out with a ratio of 0.7X and 0.15X of magnetic beads / PCR products. Tagged amplicon pools were quantified using the Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK) and pooled with equimolar concentrations into a unique library. A final clean-up was performed using a double-size selection bead purification with a ratio of 0.7X and 0.15X of magnetic beads/library. The final library was checked for size and integrity using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA) and then quantified using qPCR with the NEBNext® Library Quant Kit for Illumina® (New England Biolabs® Inc., MA, USA). Following qPCR, the library was diluted to 4 nM and 12 pM loaded on the Illumina MiSeq® with 10% PhiX using a 2 x 250 bp v2 chemistry (Illumina Inc., CA, USA).

Bioinformatics and statistical analyses

Raw sequencing data were demultiplexed using a custom Python script and subsequently analysed with metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.11 (https://github.com/HullUni-bioinformatics/metaBEAT), an in-house developed pipeline. Quality trimming, merging, chimera detection, clustering and taxonomic assignment against a custom-curated 12S reference database (Hänfling *et al.*, 2016) containing sequences for all UK freshwater fish species were performed for the taxonomic assignment using a minimum identity of 100%. metaBEAT results were summarised as the number of reads assigned to OTUs at each sample screened, from now on named as read counts. A low-frequency noise threshold of 0.001 (0.1%) was applied across the dataset to reduce the probability of false positives arising from cross-contamination or tag-jumping (De Barba *et al.* 2014; Hänfling *et al.* 2016). The threshold was applied upon the proportion of reads assigned to each fish OTU over the

total read counts calculated on a sample by sample basis. Furthermore, based on the low fish read counts found in sampling/filtration blanks and PCR negatives, an additional threshold was applied and all OTUs with less than 25 reads assigned were removed from downstream analysis.

We tested the hypothesis of no significant difference in sensitivity of qPCR and metabarcoding by comparing Arctic charr detection probabilities for individual samples between methods. Detection probabilities were then calculated for each sample and method as the number of positive PCR replicates over the number of total replicates (12 for both methods and for each of the 23 samples). A Pearson's correlation test between detection probabilities of qPCR and metabarcoding was used to evaluate the strength of the association between results of both approaches and a paired t-test was performed to check differences between the mean detection probability of the methods tested.

The package eDNAoccupancy (Dorazio and Erickson, 2017) was used to fit Bayesian, multi-scale occupancy models to estimate the conditional detection probability of Arctic charr in PCR replicates of both molecular methods (p) given the assumption of a constant probability of Arctic charr occurrence at each of the three habitats surveyed (SRN, SNTH, OF) (ψ) and a constant conditional probability of occurrence in eDNA samples (ϑ). Detections of Arctic charr were modelled as a function of the total DNA concentration (as assessed from Qubit essays) and the 12S DNA concentration (quantified using a SYBRGreen qPCR assay). Covariates were centred and scaled to have a mean of zero prior to fit the model. Estimates were computed using 11000 iterations of the Monte Carlo Markov Chain algorithm. To assess the relative importance of each covariate in relation to Arctic charr detections in PCR replicates and compare competing models, the model-selection criteria of posterior-predictive loss (PPLC) and widely applicable information (WAIC) were computed. All statistical analyses and graphics were computed in R v. 3.5.1 (R Core Team, 2018).

4.4 Results

Targeted qPCR – assay validation

The ecoPCR (probe not included) against the custom curated cytB database of UK freshwater fish species confirmed the specific amplification of Arctic charr with non-specific amplifications occurring only when more than three mismatches are allowed between primers and their binding regions. The primer-BLAST search against the full NCBI nucleotide (nt) database indicated that the chosen primer pairs would amplify other *Salvelinus* species not present in the UK, however the probe sequence, not included *in silico* is likely to improve the assay specificity. Genomic DNA from four closely related species present in the UK (brown trout, Atlantic salmon, rainbow trout and brook trout) produced no amplification when tested against the qPCR assay developed here (Appendix 3: Fig. S1). Environmental DNA extracts from freshwater bodies where Arctic charr is absent and related species present also showed no amplification. The established limit of detection of our assay was 5 copies/µL, as at this concentration, no amplification of the gBlocks® fragment occurred, whereas the limit of quantification was assessed to 50 copies/µL and, at this concentration, amplifications were not consistent across replicates (Appendix 3: Fig. S2).

Sample amplification and sequencing

None of the eDNA samples showed PCR inhibition when tested using the EXO-IPC assay and our assay with spiked Arctic charr genomic DNA as the average C_t calculated between replicates did not differ of over 0.5 C_t from the C_t of control samples with no eDNA template (Appendix 3: Fig. S3). The qPCR assay had average amplification efficiency of 90.92% (range 89.168-92.661%) and average R^2 value of 0.997 (range 0.996-0.998) for the standard curve, and no amplification occurred in negative controls.

The 23 eDNA samples used in this study amplified below the limit of quantification (see standard curve in Appendix 3: Fig. S4).

The MiSeq run generated a total of 22,818,774 sequences with an averaging sequencing depth of 72,671.25 reads per sample and an overall Phred Q30 score of 89%. 14,513,878 sequences survived the quality trimming step of which over 81% were

successfully merged. Following chimeras' sequences removal and clustering, the final number of total sequences for the library accounted to 7,621,276. Excluding control samples an average of 9,363.93 (± 9,284.07) reads per sample were assigned to fish species, and a total number of 23 fish species were identified. Total number of reads in negative controls ranged from 13 to 271 and, after the low-noise thresholds, no reads assigned to fish remained. The cleaned dataset had 17 fish assigned and seven species (silver bream, *Blicca bjoerkna;* crucian carp, *Carassius carassius*; rainbow trout, *Oncorhynchus mykiss;* nine-spined stickleback, *Pungitius pungitius;* rudd, *Scardinius erythrophtalmus*; European grayling, *Thymallus thymallus*) were removed after applying thresholds.

Detection probability of Arctic charr in qPCR and metabarcoding

Overall, metabarcoding and qPCR detections of Arctic charr were highly similar, with 66% and 61% of positive amplifications respectively (Fig. 2). For both molecular approaches the number of positive detections of Arctic charr DNA per sample ranged between a minimum of three and a maximum of 12 over 12 PCR replicates, and all samples amplified Arctic charr DNA using both methods. The paired t-test indicated that there was no significant difference between the average detection probability of the two molecular methods (t = 1.3835, df = 22, p-value = 0.1804; Fig. 3 B), and a highly significant and positive correlation was observed between Arctic charr detection probabilities of methods (r= 0.66, p < 0.000; Fig. 3 A).
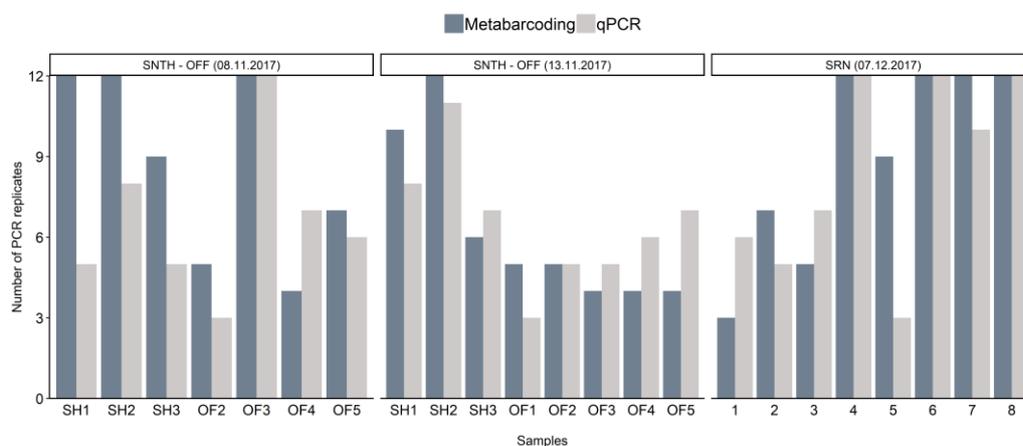
Figure 2. Barplots showing number of positive detections of Arctic charr over 12 PCR replicates for each method (qPCR and metabarcoding) across 23 eDNA samples collected at three sites (different panels) in Windermere at different sampling occasions.
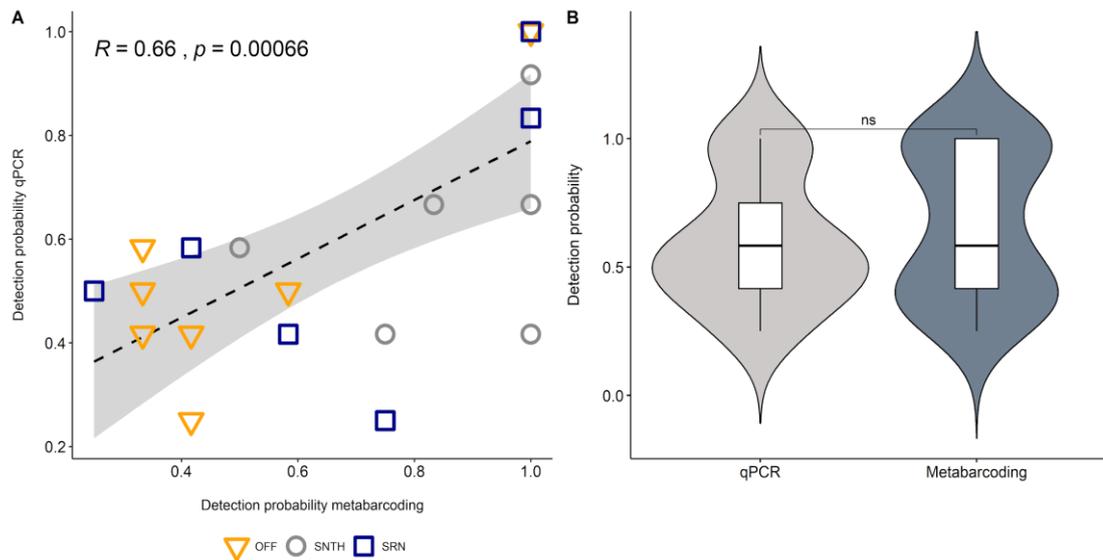
Figure 3. (**A**) Scatterplot showing association and Pearson's correlation analysis between metabarcoding and qPCR detection probabilities of Arctic charr in eDNA samples collected at different sites (SRN, SNTH and OFF) at Lake Windermere. Detection probability of each sample is calculated upon the number of positive PCR replicates over the total number of PCR replicates (*n*=12). (**B**) Violin plots showing distribution of detection probabilities across samples for different molecular methods. Difference between means of detection probability was calculated using a paired t-test.

Model-selection criteria (PPLC and WAIC) determined that 12S DNA concentration had the most significant positive effect on Arctic charr detections in metabarcoding PCR replicates (Table 1; Fig. 4A). A less significant but similar positive relationship was observed between total DNA concentration and Arctic charr detections in metabarcoding PCR replicates (Fig. 4B). Estimates of model-selection criteria additionally showed that both total DNA concentration and 12S DNA concentration significantly influence Arctic charr detection in qPCR replicates (Table 1), and positive and pronounced effects are observed for both covariates (Table 1; Fig. 4C, D). The additive occupancy models (both covariates included) estimated Arctic charr detection probability in metabarcoding PCR replicates ranging from 0.60 to 0.75 (Table 1; Appendix 3: Table S1), and Arctic charr detection probability in qPCR replicates ranging from 0.53 to 0.71 (Table 1; Appendix 3: Table S2).

Table 1. Parameter estimates (posterior medians) and model-selection criteria (PPLC and WAIC) for each of the models fitted to estimate detection probability of Arctic charr in PCR replicates for metabarcoding (upper panel) and qPCR (lower panel). Each model was fitted by running the MCMC algorithm for 11000 iterations and retaining the last 10000 per posterior summaries. Bold font indicates the lowest values of PPLC and WAIC.

Model	Occupancy in habitats (ψ)	Occupancy in sample (ϑ)	Detection in replicate (p)	PPLC	WAIC
$\Psi(.)$, $\vartheta(.)$, p(12S DNA + total DNA)	0.84	0.99	0.60-0.75	322.08	90.72
$\Psi(.)$, $\vartheta(.)$, p(12S DNA)	**0.84**	**0.98**	**0.60-0.74**	**320.82**	**87.60**
$\Psi(.)$, $\vartheta(.)$, p(total DNA)	0.84	0.83	0.64 - 0.70	329.17	89.97
$\Psi(.)$, $\vartheta(.)$, $p(.)$	0.84	0.99	0.66	329.14	87.08
$\Psi(.)$, $\vartheta(.)$, p(12S DNA+ total DNA)	**0.84**	**0.99**	**0.53-0.71**	**249.86**	**73.3**
$\Psi(.)$, $\vartheta(.)$, p(12S DNA)	**0.84**	**1**	**0.53-0.69**	**252.24**	**70.90**
$\Psi(.)$, $\vartheta(.)$, p(total DNA)	0.84	1	0.57 - 0.67	257.23	72.37
$\Psi(.)$, $\vartheta(.)$, $p(.)$	0.84	0.79	0.59	267.03	72.25
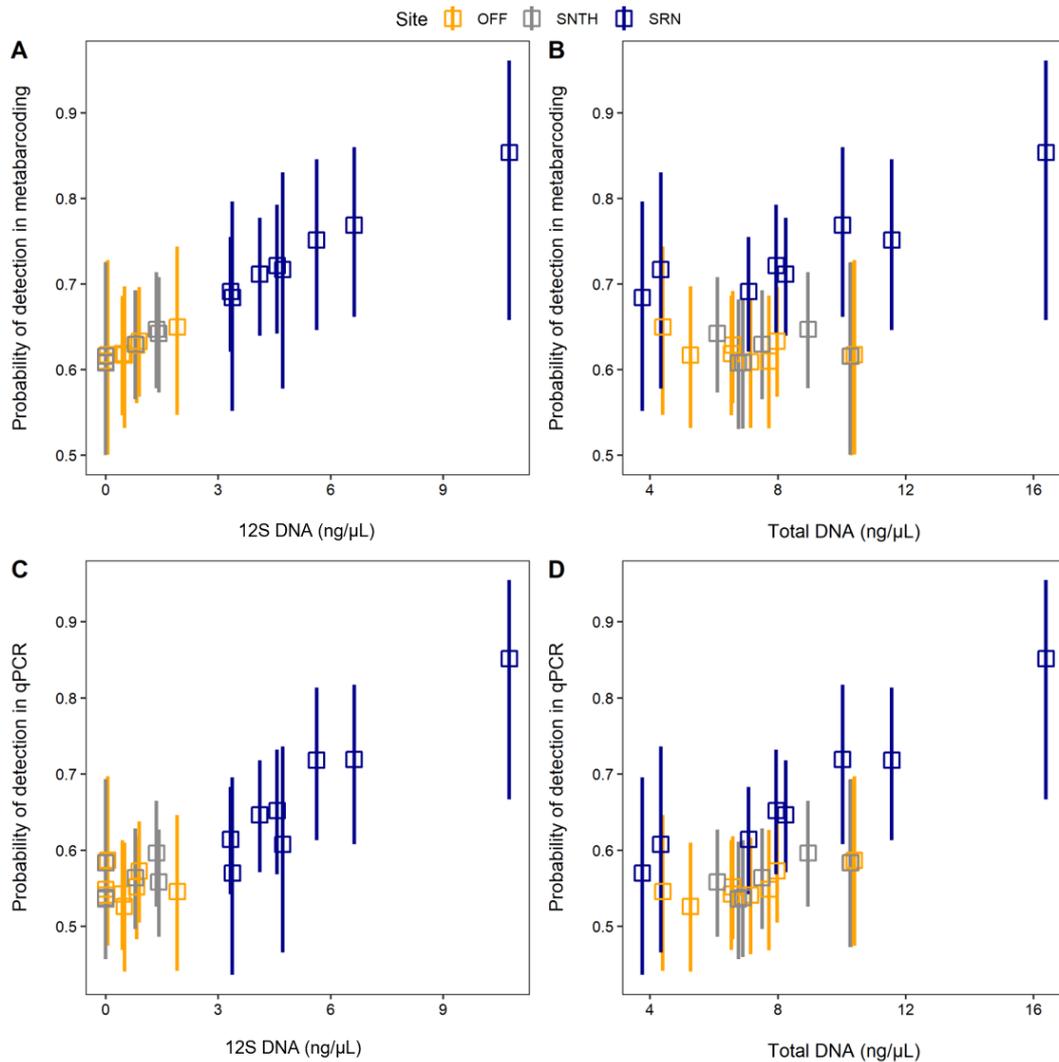
Figure 4. Estimated relationships between 12S DNA (**A**, **C**) and total DNA (**B**, **D**) concentrations and detection probabilities of Arctic charr in PCR replicates for metabarcoding (upper plots) and qPCR (lower plots). The figures show estimates of posterior medians for each eDNA sample (*n*=23) with 95 % credible intervals.

4.5 Discussion

In the present study we developed and validated a species-specific qPCR assay for the identification of Arctic charr in UK aquatic systems. We then used this targeted approach and a metabarcoding assay for UK freshwater fish identification to compare detection probability of Arctic charr between the two molecular methods using 12 PCR replicates for both molecular approaches. Overall, qPCR and metabarcoding showed congruent results with positive detections of Arctic charr across all samples analysed with both methods. Our observations additionally showed that detection probabilities

of Arctic charr in water samples did not differ significantly between eDNA methods. The lowest detection probability, as assessed by the eDNA occupancy model used in this study, was 0.53 for the qPCR assay and 0.60 for metabarcoding. The occupancy model also showed that Arctic charr detection probability in PCR replicates is influenced by the original DNA concentrations of eDNA samples, suggesting that molecular methods should account for this variable when optimising laboratory protocols. This study demonstrates that both qPCR and metabarcoding assays used in this study are similarly sensitive and accurate for surveying Arctic charr in UK freshwater habitats.

Similar detection probability of Arctic charr in eDNA metabarcoding and qPCR

Species-specific molecular approaches are often used to overcome some of the challenges associated to metabarcoding analyses as they are designed to amplify specific DNA strands within the environmental mixture (Lodge *et al.*, 2012; Laramie *et al.*, 2015; Dougherty *et al.*, 2016; Yusishen *et al*., 2020). However, over the last ten years of research in the field of eDNA metabarcoding for biomonitoring applications, the optimisation of metabarcoding approaches together with the fast-advancing sequencing technology have nowadays minimised the sources of false negative errors described in the past, such as low sequencing depth, inefficient sequencing platforms and bias of broad-range primers with the preferential amplification of certain taxa in a mixed DNA template (Deagle *et al.*, 2013; Smith and Peay, 2014; Schenekar *et al.*, 2020). In the present study, we used a fish eDNA metabarcoding protocol that has been optimised through the years for the efficient and accurate detection of UK freshwater fish species. The 12S broad-range primers used in this study perfectly match the DNA binding regions of UK freshwater fish species DNA and they have successfully detected Arctic charr in a number of previous eDNA studies (Chapter 3; Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019; Li *et al.*, 2019a). Although the 12S primers used here have no mismatches for the Arctic charr DNA annealing regions, primers biases in eDNA metabarcoding can also occur for the preferential amplification of more abundant DNA strands within the mixed DNA template during the amplification (Bylemans *et al.*, 2018; Bylemans *et al.*, 2019; Elbrecht and Leese, 2015). In addition, PCR stochasticity with the random amplification preference for specific

DNA strands might also add another level of uncertainty (Kebschull and Zador, 2015). These two latter sources of false-negative errors in metabarcoding studies are likely to affect low-abundant DNA sequences such as Arctic charr DNA template in a number of water samples used for this study (Fig. 1). To circumvent this stochastic effect, we used a high number of PCR replicates (n=12) for both molecular approaches (qPCR and metabarcoding) as suggested from previous studies (Schultz and Lance; 2015; Ficetola *et al.*, 2015; Piggot *et al.*, 2016) and we found that all eDNA samples analysed efficiently detected Arctic charr DNA with both methods, including samples that failed to detect this species in previous eDNA metabarcoding analyses (Fig. 1,2). Using the same level of replication between targeted and whole-community eDNA approaches we additionally observed that detection probability of Arctic charr was similar between the methods (t = 1.3835, df = 22, p-value = 0.1804; Figure 3B) and significantly correlated (r= 0.66, p < 0.000; Figure 3A). This result is in contrast with previous studies which showed higher detections of target species in species-specific molecular approaches as opposed to metabarcoding (Lacoursière-Roussel *et al.*, 2016c; Harper *et al.*, 2018; Bylemans *et al.*, 2019; Wood *et al.*, 2019). Previous comparisons, however, used different levels of replication and a higher number of PCR replicates in qPCR compared to metabarcoding (Harper *et al.*, 2018) or qPCR primers amplifying shorter DNA regions compared to those amplified using broad-range metabarcoding primers (Lacoursière-Roussel *et al.*, 2016c) or using metabarcoding primers with mismatches for the target species DNA binding regions (Bylemans *et al.*, 2019; Wood *et al.*, 2019). The use of primers amplifying shorter DNA targets is advantageous for eDNA studies as shorter regions have higher chances of amplification compared to the larger ones due to the degraded nature of the environmental DNA (Hänfling *et al.*, 2016). Primers for species-specific assays are usually designed to amplify very small target DNA regions (typically between 60 and 150 bp; Bustin and Huggett 2017) and maximise detection of a single species, whereas metabarcoding primers aim to achieve the broadest taxonomic coverage within a group of species and have, therefore, larger DNA target regions. The different key objectives of the two molecular approaches (target species vs. community approach) is arguably the reason of the different sensitivity of both molecular approaches toward target species as both methods are optimised to address different ecological questions. In our study there was also a difference between

amplified regions of both molecular approaches (39 bp), and in our assays the qPCR target DNA region was longer compared to the metabarcoding one (148 bp vs. 106), however this difference did not introduce variation in detection probability of Arctic charr as we found no significant difference between the sensitivity of the methods.

To enhance detection probability of Arctic charr in metabarcoding we have used PCR replicates which were sequenced separately, thus increasing sequencing depth and the sensitivity of the whole-community method, making this approach more comparable to the qPCR (Grey *et al.*, 2018; Kelly *et al.*, 2014). This metabarcoding approach is of difficult implementation for large-scale monitoring studies due to high costs for large number of samples. For this reason, PCR replicates are usually pooled to reduce the sequencing effort in eDNA metabarcoding studies (Hänfling *et al.*, 2016; Harper *et al.*, 2018; McClenaghan *et al.*, 2020). A trade-off between level of replications and costs would certainly improve the sensitivity of metabarcoding methods with more accurate detections of low-abundant species and decreased chances of false-negatives.

Original DNA concentration influence the detection of target Arctic charr in eDNA samples

The occupancy models suggest that detection probability of Arctic charr in PCR replicates are influenced by the 12S DNA concentration and total DNA concentration in eDNA samples, with original 12S DNA concentration having a significant effect on Arctic charr detections in metabarcoding PCR replicates. Piggot (2016) already found that the original sample eDNA yield influences target species detection probability in environmental samples. These results may be integrated into future eDNA-based monitoring programmes where the number of PCR replicates can be adjusted based on a previous evaluation of the total and group-specific DNA yield. A high level of replication for samples with high starting genetic material might be unnecessary and might be avoided in order to optimise costs, time and resources. In addition, the number of PCR replicates should also be adjusted based on other biotic or abiotic factors as well as on the aim of the study. For example, the eDNA samples used in the present study were collected during the Arctic charr spawning season, hence, a lower level of technical replication would be sufficient to detect the species at those

84

spawning sites (Fig. 1; Chapter 3), whereas higher number of PCR replicates might be required if sampling the same sites outside the breeding period (Chapter 3). Abiotic factors such as water mixing might also influence detection of the study species in PCR replicates as, for instance, stratification limit the vertical transport of DNA in water, and a higher number of PCR replicates would be beneficial for the detection of deep water species in shore eDNA samples (Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019; Zhang *et al.*, 2020).

4.6 Conclusion

Our results demonstrated that the qPCR assay developed in this study efficiently detects Arctic charr in eDNA samples from UK aquatic systems, and that the eDNA metabarcoding protocol used is not less sensitive than qPCR when both methods are compared using the same technical replication effort. Although a higher level of PCR replication in metabarcoding approaches would be desirable to enhance low-abundant species detection probability, this would be unmanageable for broad-scale or long-term monitoring surveys. However, as demonstrated in the present study, a number of technical, biological and environmental factors can be used to adequately adjust the level of replication for the effective detection of the species of interest. Species detection in eDNA-based studies is always a balance between number of eDNA samples collected, volumes of water filtered and PCR replicates, and the level of replication used in each of these steps should be based on the aim of the study, target species, habitats surveyed and resources available (Chapter 2; Li *et al.*, 2018; Doi *et al.*, 2019). Using whole-community approaches for evaluating the status of a target species is beneficial because these holistic approaches have the additional advantage of simultaneously address multiple ecological questions such as, the detection of non-native species (Blackman *et al.*, 2017; Klymus *et al.*, 2017), spatio-temporal changes of species composition (Lawson Handley *et al.*, 2019; Harper *et al.*, 2020), estimates of species abundance (Chapter 2; Hänfling *et al.*, 2016; Li *et al.*, 2019a), species spawning activity (Chapter 3), and identification of intra-specific variation (Elbrecht *et al.*, 2018). All these additional advantages might hold essential information to understand the drivers contributing to a change in the status of a target species within its ecosystem

and, therefore, the use of eDNA metabarcoding for biomonitoring might be much more informative for the species and ecosystem conservation.

Chapter 5

Environmental DNA (eDNA) metabarcoding provides accurate information on distribution and abundance of UK priority conservation fish

Searching vendace DNA at Loch Skeen. Photo credit: Dr. Lawson Handley L., used with permission.

5.1 Abstract

Anthropogenic changes are threatening native fish populations and their freshwater habitats through eutrophication, habitat fragmentation, climate change and species introductions. To combat the decline of endangered fish populations, a variety of management actions can be initiated. These include the monitoring of the distribution and abundance of native populations, habitat restoration and *in-situ* management (such as reducing exploitation), and the translocation of fish to refuge sites. Current methods to assess the status of native and translocated populations of endangered fish (e.g. gill netting or hydroacoustics) are limited or inefficient as they are not universally applicable, costly in terms of time and resources, and can result in the

mortality of individuals. Here we use environmental DNA (eDNA) metabarcoding to monitor the status of three threatened salmonid species in UK lentic systems. Existing fish eDNA metabarcoding data from 101 UK lakes were used to evaluate the accuracy of this approach to provide information on the distribution and abundance of Arctic charr and coregonids (vendace and European whitefish). In this study, estimates of abundance from site occupancy of eDNA metabarcoding were scaled to abundance scores (from 0 to 5, absent to dominant) and compared to historical estimates of abundance. Our results show that eDNA metabarcoding reflects the species expected distribution, and that abundance estimates from historical data and eDNA site occupancy were positively associated. We conclude that eDNA metabarcoding results can assist conservation actions through the identification of sites with potentially endangered populations, and be used as a non-invasive tool for evaluating the status of newly translocated populations in refuge sites.

5.2 Introduction

In the ongoing Anthropocene and in the face of a sixth mass extinction, the goal of conservation efforts is to attenuate the loss of biodiversity and preserve ecosystem services worldwide (McCallum, 2015; Schwartz *et al.*, 2017). One of the key aspects of conservation programmes is the ability to assess spatial distribution and trends in abundance of rare and threatened species, and to evaluate the success of management actions (Maxwell and Jennings 2005; Joseph *et al.*, 2006; Poos *et al.*, 2007). Beyond financial, human and technical resource limitations, the ability to locate rare or threatened fish species in aquatic systems is problematic due to low visibility in turbid or deep waters and habitat complexity (Cooke *et al.*, 2016). Moreover, many freshwater fish species are elusive and highly mobile, and sampling methods for population assessment are limited or inefficient (Maxwell and Jennings 2005). Failure to detect the presence of species when they are actually present (false negatives) is an important issue in threatened species management and can bias population assessment leading to ineffective management actions (Delaney and Leung, 2010; Cooke *et al.*, 2016).

In the UK, three lacustrine salmonid species, Arctic charr (*Salvelinus alpinus* L.),
European whitefish (*Coregonus lavaretus* L.), and vendace (*Coregonus albula* L.) are
considered vulnerable as they have a restricted natural distribution and are threatened
by several anthropogenic pressures including climate change, eutrophication,
acidification, and the introduction of non-native species (Winfield *et al.*, 2006;
Maitland *et al.*, 2007; Winfield *et al.*, 2008a; Winfield *et al*., 2010). Understanding and
promptly tackling these issues is essential to safeguard these species over their
geographical distribution. For example, the early detection of newly introduced
competitors (e.g. roach, *Rutilus rutilus*) would determine effective responses and avoid
the decline of these threatened populations at their native sites (Winfield *et al.*, 2006;
Winfield *et al.*, 2008a). Monitoring environmental parameters (e.g. water chemistry
and temperature) is also paramount to evaluate changes that could affect threatened
populations (Winfield *et al.*, 2010; Graham and Harrod, 2009; Kelly *et al*., 2020). The
assessment of drivers influencing these threatened species distribution and abundance
would benefit the effective management of these species across their limited
distribution range.

These three salmonid species inhabit deep oligotrophic or mesotrophic lakes with
suitable spawning substrate and few predators, and these requirements limit the
distribution of these species to a few water bodies in the UK (Maitland and Lyle, 2013).

Historically Arctic charr have been recorded in 295 Scottish lochs (Adams and
Maitland, 2018) and many of these sites have not been surveyed so the current status
of the populations is unknown and at least 12 Arctic charr populations are considered
extinct (Maitland *et al.*, 2007). In England, eight populations of Arctic charr are
restricted to the Lake District (Cumbria) and concerns about their conservation status
exist over almost all their native range. Only four populations have ever been recorded
in Wales and, of these, one is now extinct.

Vendace is the rarest freshwater fish in the UK, having only ever been recorded in
historical times from four locations, two in the Lake District (Bassenthwaite Lake and
Derwent Water) and two in Scotland (Castle Loch and Mill Loch). In Scotland, habitat
degradation and the introduction on non-native species led to the loss of both native

populations. The Bassenthwaite population was rediscovered in 2013 when, after 12 years of lack of records, a single individual was caught during a netting survey and two adults were recorded the following year (Winfield *et al.*, 2017).

European whitefish is the most widely distributed coregonid species in the UK, but is still considered rare because it is only found in seven native sites in the English Lake District, two sites in Scotland, and one in Wales (Winfield *et al.*, 2013).

In the UK, Arctic charr, vendace and European whitefish populations are listed as Priority Species within the UK Biodiversity Action Plan (www.ukbap.org.uk), and some populations are formally protected in sites designated for nature conservation (i.e. Sites of Special Scientific Interest, SSSI) (Maitland and Lyle, 2013). Both European whitefish and vendace also receive strict protection under Schedule 5 of the Wildlife and Countryside Act 1981 wherever they occur. Populations of all three species which occur within protected sites, or in waterbodies to which they have been established as ark or refuge sites (see Maitland and Lyle, 2013; Adams *et al.*, 2013), are periodically monitored according to the reporting requirements of the EC Habitats Directive (European Communities, 1992). Standard sampling methods for monitoring Arctic charr and coregonids include the use of quantitative hydroacoustics to determine population abundance, and targeted Norden-style gill net surveys to support hydroacoustic data and to obtain specimens used to determine other biological characteristics, such as growth and condition (Bean, 2003a, 2003b; Bean, Hall and Thomas, 2015).

Conventional fish-sampling methods, such as gill netting, are often destructive and may not be desirable when sampling species of high conservation value. Alternative methods of counting fish in lakes which are not destructive, such as the use of hydroacoustics, can be limited by lake bathymetry and depth, and lack the ability to differentiate between species when not accompanied with other sampling methodologies (Winfield *et al.*, 2009). Consequently, a combination of sampling methods, such as hydroacoustics and gill netting, is frequently used to monitor freshwater fish populations in standing waters. Based on a number of previous studies, environmental DNA (eDNA) is a promising complementary tool to assess the status of threatened fish species in sensitive sites (Laramie *et al.*, 2015; Sigsgaard *et al.*, 2015;

Kamoroff and Goldberg, 2018). For example, eDNA analysis has been used to monitor the success of reintroductions of locally extinct fish species (Riaz *et al.*, 2019; Hempel *et al.*, 2020) and the dispersal of non-native fish species (Mahon *et al.*, 2013; Takahara *et al.*, 2013; Hinlo *et al.*, 2017). Recent studies have also promoted the potential use of eDNA metabarcoding as a tool for assessment of freshwaters ecosystems for the EU Water Framework Directive (Hering *et al.*, 2018; Li *et al.*, 2019; Pont *et al.*, 2019).

Here, we evaluate distribution and abundance estimates of Arctic charr, vendace and European whitefish using eDNA metabarcoding data from an existing collection of 101 UK freshwater bodies which include protected sites (i.e. SSSIs), translocation (ark or refuge) sites, and sites where these species are considered to be extinct or severely endangered. Fish metabarcoding data were used in this chapter to evaluate: i) the agreement between historical data (based on conventional surveys and expert opinion) and eDNA metabarcoding data on distribution and abundance of Arctic charr and coregonids; ii) the status of these rare species at five translocation sites; iii) biological and environmental factors driving the current distribution of Arctic charr and coregonids in the UK.

5.3 Materials and methods

Study sites

The list of 101 UK freshwater bodies sampled from which the data in this study was obtained includes sites surveyed within a number of joint projects with the Scottish Environment Protection Agency (SEPA), the Environment Agency (EA) and the Food and Environment Research Agency (FERA), aiming to the implementation of fish eDNA metabarcoding into the WFD or similar monitoring programmes (Willby *et al*., 2019). In five additional translocation sites (Loch Earn, Loch Skeen, Grimwith Reservoir, Kielder Water, Sprinkling Tarn), the eDNA sampling occurred together with gill netting surveys (with the exception of Sprinkling Tarn), and these refuge sites were surveyed to assess the status of the target species after translocation. Grimwith Reservoir and Kielder Water are Arctic charr refuge sites. In Grimwith Reservoir young specimens from the autumn-spawning population of Windermere South basin were introduced in 1989, 1990 and 1991 (Maitland *et al.*, 2006). Introduction of Arctic charr into Kielder Water is

more recent (2013), and 10,000 juveniles from Ennerdale Water have been stocked in this reservoir. To monitor establishment of Arctic charr in Grimwith Reservoir, gill netting surveys were carried out after translocation, indicating reasonable initial survival and growth up to 1998, but more recent surveys failed to record any Arctic charr specimens and the current status of the population had been listed as 'unknown' (Maitland et al., 2006). Similarly, there are uncertainties regarding the establishment of Arctic charr in Kielder Water, where a lack of resources has prevented commissioning of surveys to assess the success of the translocation programme.

Loch Skeen, Loch Earn and Sprinkling Tarn are refuge sites for vendace. In Loch Skeen the translocation of fry and eyed eggs between 1997 and 1999 has resulted in the rapid establishment of a self-sustaining population (Maitland and Lyle; 2013; Adams et al., 2014). Conversely, the establishment of vendace in Loch Earn after the translocation of fry from Bassenthwaite Lake in 1989, appears to have been much slower, and the introduction was considered to be a failure until a single specimen was caught by an angler in 2005 (Lyle and Adams, 2017). Translocation of vendace eggs from Derwent Water to Sprinkling Tarn occurred in 2005 (Lyle, Maitland and Winfield, 2006), but the success of the introduction remains unknown as the site has been poorly monitored since the translocation.

Overall, our dataset comprises 44 Scottish lochs, 49 English lakes (21 located in Cumbria), and 8 Welsh lakes (Fig. 1; Appendix 4: Table S1). Some sites are listed as SSSI for Arctic charr such as Loch Doon, Loch Insh, and Loch Eck in Scotland, Windermere, Wastwater, Crummock Water, Ennerdale Water and Buttermere in the English Lake District (Cumbria), and Llyn Padarn and Llyn Cwellyn in Wales (Appendix 4: Table S1). Loch Doon and Windermere Arctic charr populations are also monitored because they are threatened by acidification and eutrophication respectively (Maitland et al., 2006; Maitland and Lyle, 2013). Additional sites known to hold vulnerable Arctic charr populations included in the dataset are those from Loch Lee, Loch Awe and Coniston Water (Winfield et al., 2004; Maitland et al., 2006). Several other sites, where Arctic charr are now considered to be extinct (Loch Grannoch, Loch Achray, Loch Eilt, Loch Katrine, Loch Affric, Ullswater, Loweswater; Appendix 4: Table S1), were also included in this study.

The dataset additionally contains six water bodies within the limited range of distribution of coregonids such as Bassenthwaite Lake and Derwent Water in the English Lake District where native vendace populations occur, and Loch Lomond, Loch Eck, Ullswater and Brothers Water with native populations of European whitefish.

Sample collection and processing

A standard approach for sample collection, processing and metabarcoding workflow was followed across the sites surveyed following general guidelines as described in Hänfling *et al.*, 2016 and in the previous chapters.

Environmental DNA sample collection occurred in the autumn/winter season between 2016 and 2019. 2 L water bottles (Gosselin™ HDPE plastic bottles; Fisher Scientific UK Ltd, UK) were used for the collection of shore and offshore samples (when applicable; see Appendix 4: Table S1 for details on sample number). Collection blanks (2 L molecular grade water) were used at each site and water samples were vacuum-filtered within 24 hours from collection. All samples, collection and filtration blanks, were filtered using 0.45 μm mixed cellulose ester filters (GE Healthcare), and filters were stored at -20 °C until DNA extraction. The mu-DNA protocol for water samples was used for DNA extractions (Sellers *et al.*, 2018). Aliquots of the final purified DNA for each sample were measured with a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) to assess yield and purity and DNA extracts were then frozen at -20 °C until PCR amplification.

Sequencing libraries were prepared using vertebrate-specific primers targeting a 106 bp fragment of the mitochondrial 12S ribosomal RNA (rRNA) region in fish (Riaz *et al.*, 2011; Kelly *et al.*, 2014) following a double-tagging approach based on two PCRs as described in the previous chapters. The first PCRs were performed in triplicate and samples from individual lakes were pooled in sub-libraries. Sub-libraries were cleaned using a double-size selection magnetic beads protocol (Bronner *et al.*, 2013) and Illumina tags were added in the 2nd PCRs. A second double-size selection bead purification was performed and then sub-libraries were pooled at approximately equal concentration. The final libraries were quantified by qPCR, checked using the Agilent

2200 TapeStation for size and integrity, and finally sequenced on an Illumina MiSeq®
using 2 x 300 bp V3 chemistry (Illumina Inc., CA, USA).

Bioinformatics and statistical analyses

Raw Illumina sequences were demultiplexed using a custom Python script and then
processed using metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.11
(https://github.com/HullUni-bioinformatics/metaBEAT). The pipeline performed
trimming, quality control (filtering high quality reads and chimera detection) and
comparison with our custom reference database for UK fish (Hänfling *et al.*, 2016).

The data produced by metaBEAT were summarised as the number of sequence reads
per fish species detected in each sample. Downstream analyses were performed in R
v3.5.1 (R Core Team, 2018) and graphs plotted using *ggplot2* (Wickham, 2016). The
total read counts per sample were used to calculate the proportional read counts for
each fish species. We then applied a false positive sequence threshold of 0.001 (0.1%)
to remove taxonomic assignments that may have resulted from contamination during
library preparation or sequencing (De Barba *et al.*, 2014; Hänfling *et al.*, 2016; Port *et
al.*, 2016). For downstream analyses, the data was condensed to the proportion of
sampling sites in which a given species was detected (i.e. site occupancy).

Following a similar approach as in Li *et al.*, (2019a), we calculated abundances
estimates for Arctic charr and coregonids and placed them into abundance categories
on a DAFOR scale (5 = Dominant; 4 = Abundant; 3 = Frequent; 2 = Occasional; 1 = Rare;
0 = Absent; Table 1). Estimates from historical data were inferred independently
through expert assessment (C.W.B.; I.J.W. and C.E.A.) based on a collection of past and
recent outcomes from conventional surveys (when applicable) and historical
knowledge of the species distribution. For refuge sites with no or debated data about
establishment after species translocation (i.e. Arctic charr in Kielder Water and
Grimwith Reservoir and vendace in Sprinkling Tarn) we assessed populations as rare at
those locations (assigned abundance score equal to 1). DAFOR scale abundance scores
from eDNA site occupancy were calculated as described in Table 1.

To compare historical data and eDNA metabarcoding data on Arctic charr and coregonids abundance score we used Bland-Altman plots for visualisation and Spearman's rank correlation analysis. Maps to visualise agreement between methods at the sites surveyed were built using the *rnaturalearth* package (South, 2017). In the maps, strong agreement of abundance scores was assigned when the abundance scores between methods were equal or differ of one score on the abundance scale (e.g. 1\2, 2\3 etc. eDNA abundance\historical data abundance or *vice versa*). Moderate agreement was assigned when the difference between methods on estimated abundance scores was two, and low agreement was assigned when the difference was over two on the abundance scale.

Fish eDNA metabarcoding data were then used to evaluate biotic (lakes fish species composition) and abiotic factors (lake-specific and catchment-specific parameters) influencing distribution and abundance of Arctic charr and coregonids species. Fish species detections from eDNA metabarcoding across the dataset were used to investigate significant species associations using the R package *cooccur* (Griffith *et al.*, 2016) which evaluates if species are negatively, positively or randomly associated with one another based on the probabilistic model of species co-occurrence from Veech, 2013 using a pairwise approach. This probabilistic model measures the observed species co-occurrence (number of sites where two species co-occur) and compare the observed co-occurrence to the expected co-occurrence which is the product of the two species' probability of occurrence multiplied by the number of sampling sites (Veech, 2013; Griffith *et al.*, 2016). The significance of the associations is then obtained calculating the probability that the observed frequency of co-occurrence is significantly large than expected (positive associations), significantly small than expected (negative associations) or not significantly different from the expected (random associations) (Veech, 2013; Griffith *et al.*, 2016).

A generalized linear model (GLM) using quasi-binomial family and logit link was used to test the influence of several abiotic parameters on Arctic charr relative abundance. The number of eDNA detections of coregonids across the lakes surveyed was too limited for such an analysis. Arctic charr site occupancy was used as a dependent variable and environmental parameters selected from the online database UK Lakes Portal

(https://eip.ceh.ac.uk/apps/lakes/index.html) as independent variable. The initially selected variables included nine lake-specific parameters (i.e. perimeter length [km]; fetch of wind distance [km] which measures the distance wind travels over water before meeting an obstacle, considered here as a measure of water mixing; distance to the sea [km]; water volume [m^3]; surface area [ha]; altitude [m]; mean depth [m]; mean alkalinity [µEq/L]; morpho-edaphic index [MEI; Alkalinity/mean depth] which is an index of productivity), and eight catchment specific parameter (i.e. catchment area [ha]; land cover [agricultural, coniferous and urban in percentage]; percentage of the lake's hydrological catchment area covered by lakes or ponds; mean slope [degrees]; count of river obstacles/barriers). Prior to GLM analysis, we checked factors for collinearity using Pearson's correlations between each pair of explanatory variables. For co-varying variables with a correlation coefficient higher than 0.35 only the most ecologically significant variable of the correlated pair was retained. Model selection was performed using the *dredge* function implemented in the *MuMIn* package (Barton, 2020). The best models were selected using *quasi*-AIC, a modification of the Akaike's Information Criterion, for *quasi*-family models.

Table 1. Criteria for the assignment of abundance scores based on: i) historical data from conventional methods and expert opinion; ii) site occupancy from eDNA metabarcoding data. Abundance scores were assigned to Arctic charr and coregonids at each location surveyed (see Appendix 4: Table S1 for full list of lakes).

DAFOR categories from historical data	Site occupancy from eDNA metabarcoding	Assigned abundance score
Absent	SO = 0	0
Rare	0 ≤ SO < 0.2	1
Occasional	0.2 ≤ SO < 0.4	2
Frequent	0.4 ≤ SO < 0.6	3
Abundant	0.6 ≤ SO < 0.8	4
Dominant	0.8 ≤ SO ≤ 1.0	5

5.4 Results

Agreement between expected species distribution/abundance and eDNA

Environmental DNA metabarcoding detected Artic charr and coregonids in all freshwater bodies where the species were expected to occur based on historical data, with the exception of the refuge site of Grimwith Reservoir where Arctic charr DNA was not detected (Fig. 1; Appendix 4: Table S1). Out of the 101 water bodies surveyed within this study, Arctic charr had been historically recorded in 27 lakes and coregonids in nine lakes (Fig. 1A; Appendix 4: Table S1). Arctic charr eDNA was found in 26 lakes where the species is known to occur and in four additional lakes: Loch Lomond and Loch Lubhair, Sprinkling Tarn and Lock Skeen (note the latter two are vendace refuge sites - Fig.1A, Appendix 4: Table S1). Coregonids were detected in two additional lakes; Loch Awe and Windermere (Fig.1B; Appendix 4: Table S1).

At 19 out of the 26 sites where Arctic charr was detected both historically and with eDNA, there was a strong agreement between estimated abundance scores assigned from eDNA site occupancy and historical data (Fig. 1A; Appendix 4: Table S1). Five of the remaining seven sites showed moderate agreement between abundance scores, and two sites had low agreement (Fig. 1A; Appendix 4: Table S1).

Overall, coregonids occur at 11 sites surveyed according to historical knowledge of the species distribution and eDNA data. Estimated abundance scores assigned to coregonids from both historical data and eDNA site occupancy showed strong agreement at five sites, moderate agreement at one site and low agreement at two sites (Fig. 1B; Appendix 4: Table S1).
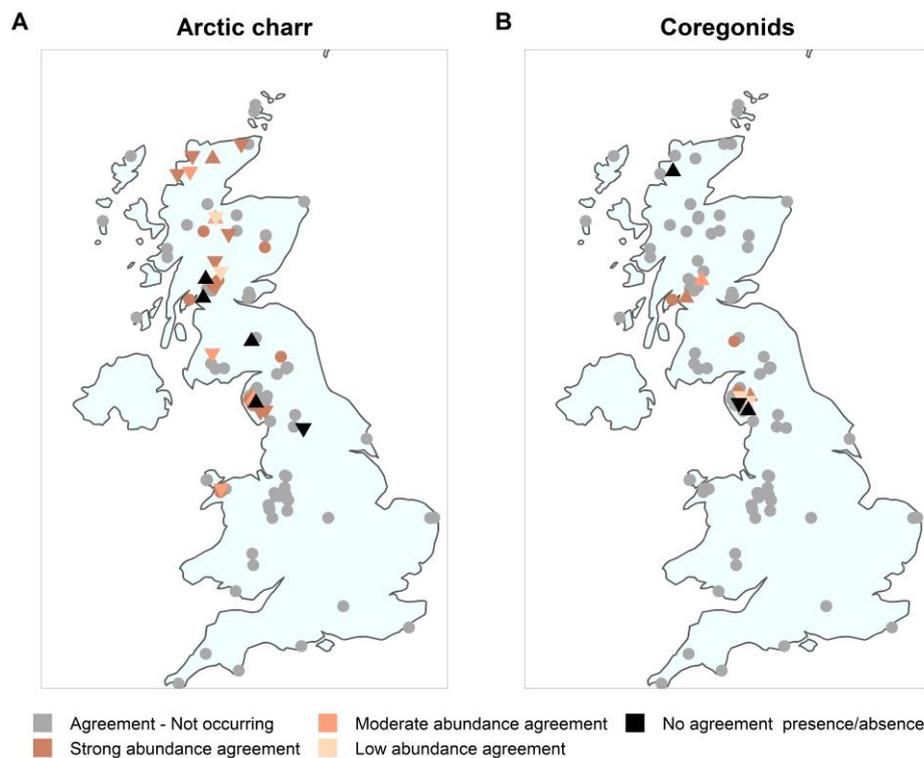
Figure 1. Map showing different levels of agreement between estimated abundance scores assigned by historical data and site occupancy from eDNA metabarcoding for Arctic charr (**A**) and coregonids (**B**) in 101 UK lakes. Triangles pointing downward indicate that historical data abundance scores are higher than abundance scores from eDNA metabarcoding, and *vice versa* for triangles pointing upward. A circle was used for lakes with equal abundance estimates between methods.

Across the entire dataset and for both Arctic charr and coregonids, a good agreement between estimated abundance scores of historical data and eDNA data was found (Fig. 2A). In the Bland-Altman plot over 95% of the data points lay between the limits of agreement and no systematic bias between methods for abundance scores calculation was observed (Fig. 2A). A total of 13 abundance scores fall outside the limits of agreement and, in seven cases, historical data overestimated species abundance compared to eDNA-assessed abundance scores (above the upper limit, Fig. 2A; Fig. 1), whereas in six cases eDNA data generated lower estimates of species abundance (below the lower limit, Fig. 2A; Fig. 1). Lastly, positive and significant Spearman's associations were found between estimated abundance scores from eDNA and historical data for both Arctic charr ($r = 0.46$, $p < 0.0089$; Fig. 2B) and coregonids ($r = 0.73$, $p = 0.011$; Fig. 2B).

A

B

$R = 0.73$, $p = 0.011$
$R = 0.46$, $p = 0.0089$

Salvelinus alpinus   Coregonus spp.

Salvelinus alpinus   Coregonus spp.
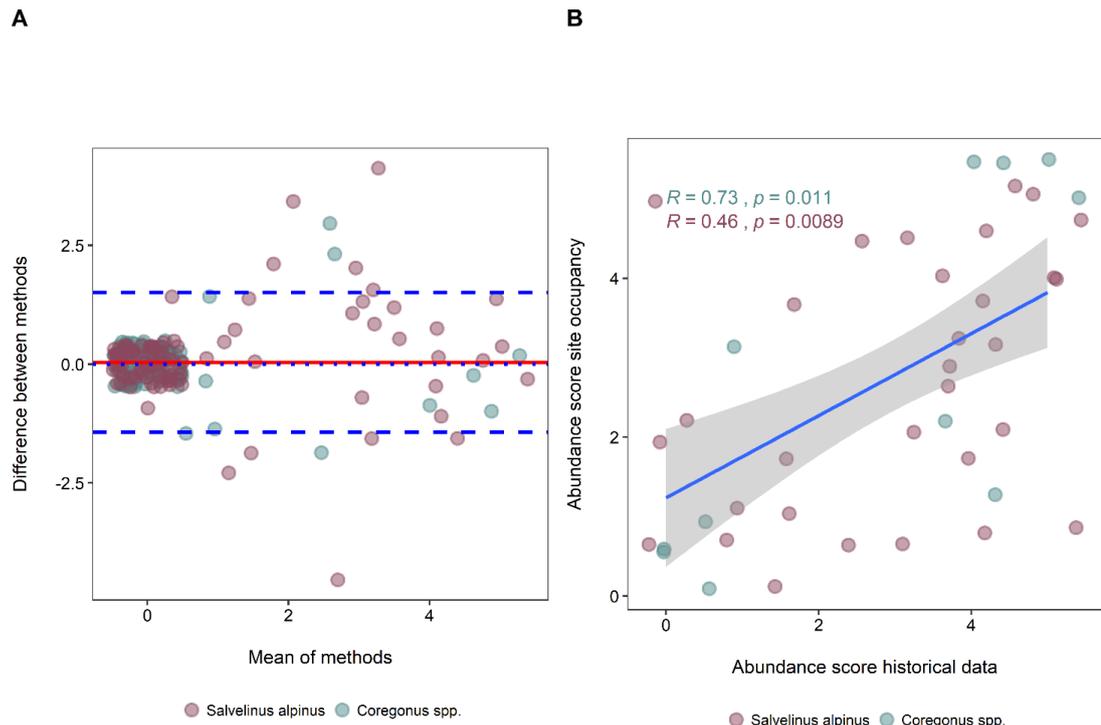
Figure 2. (**A**) Bland-Altman plot (difference plot) describing agreement between abundance scores (DAFOR scale 0 to 5) assessed by historical data (method 1) and eDNA site occupancy (method 2). In the figure the difference between the two paired measurements (for each location and taxa) is plotted against their mean. The red solid line represents the mean difference and the blue dashed lines show the limits of agreement (± 1.96 SD of the mean difference). (**B**) Scatterplot showing association between abundance scores of historical data and eDNA metabarcoding site occupancy and Spearman's correlation outputs. The 0,0 sites were not included in the statistical analysis.

Species detections at refuge sites – conventional methods versus eDNA metabarcoding

Environmental DNA metabarcoding detected more fish species than conventional methods (gill-nets or fyke-nets) at all five refuge sites surveyed (Fig. 3). Overall, a maximum of three fish species were detected using conventional methods, whereas eDNA detected between a minimum of three species in Grimwith Reservoir to a maximum of 12 species detected in Loch Earn (Fig. 3).

Neither eDNA metabarcoding nor netting methods detected Arctic charr at the refuge site of Grimwith Reservoir. Arctic charr eDNA was found in Kielder Water where

conventional surveys failed to detect this species (Fig. 3), but the species was detected in only one of 11 samples (abundance score = 1; Appendix 4: Table S1) and this result was confirmed using a targeted qPCR assay for Arctic charr as described in Chapter 4 (Appendix 4: Fig. S2). Positive and consistent detections of Arctic charr eDNA were additionally found in Sprinkling Tarn (vendace refuge site), and detections occurred at eight sites over a total of ten sites surveyed assessing Arctic charr as a dominant species within tarn (abundance score = 5; Appendix 4: Table S1; Fig. 3).

Vendace was detected at the species translocation sites of Loch Earn and Loch Skeen with both conventional and molecular methods, whereas vendace was not found in Sprinkling Tarn using eDNA metabarcoding (recent netting data are not available for this site) (Fig. 3). At Loch Skeen, estimated abundance scores assigned to vendace using conventional methods and eDNA site occupancy were identical (abundance score = 5) and evaluated the species as dominant within the loch (Fig. 1B; Appendix 4: Table S1). Conversely, vendace abundance in Loch Earn differed between assessment methods, with a higher abundance score from eDNA site occupancy (abundance score = 3; Fig. 1B; Appendix 4: Table S1) compared to conventional methods (abundance score = 1; Fig. 1B; Appendix 4: Table S1).
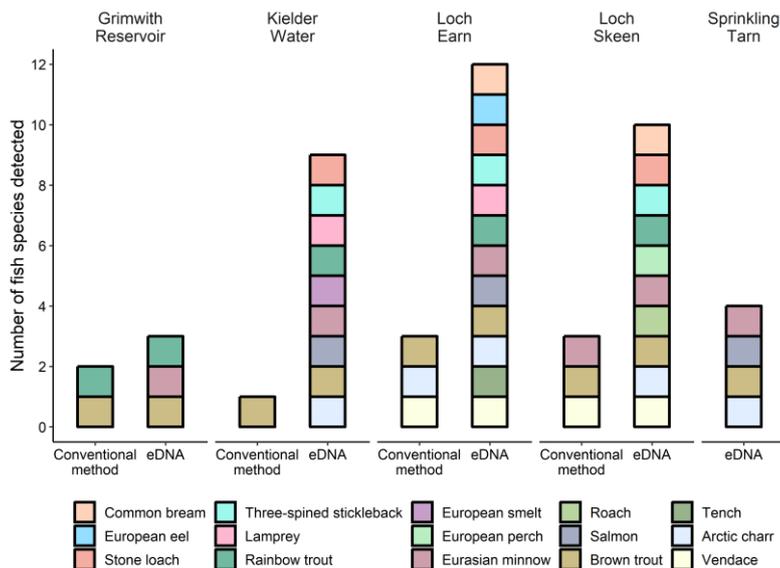

Figure 3. Bar plots of fish species detected with conventional methods (gill-nets or fyke-nets) and eDNA metabarcoding at five refuge sites surveyed. To note: results of conventional methods for Sprinkling Tarn are not available as this water body has not been surveyed using conventional approaches after vendace introduction.

Biotic and abiotic factors influencing priority conservation species distribution and abundance in UK freshwaters

Cooccur analysis found positive associations between Arctic charr and brown trout (*Salmo trutta*), Atlantic salmon (*Salmo salar*), European river lamprey (*Lampetra fluviatilis*), and Eurasian minnow (*Phoxinus phoxinus*; Fig. 4; Appendix 4: Table S2). Negative associations were found between Arctic charr and roach (*R. rutilus*), tench (*Tinca tinca*), pike (*Esox lucius*), European perch (*Perca fluviatilis*), rudd (*Scardinius erythrophthalmus*), common carp (*Cyprinus carpio*) and common bleak (*Alburnus alburnus*; Fig. 4; Appendix 4: Table S2). Coregonids were positively associated with stone loach (*Barbatula barbatula*), European river lamprey, rainbow trout (*Oncorhynchus mykiss*), nine-spined stickleback (*Pungitius pungitius*), three-spined stickleback (*Gasterosteus aculeatus*), Eurasian minnow, Atlantic salmon and Arctic charr (Fig. 4; Appendix 4: Table S2). No other significant associations were found between coregonids and other fish species.

Five independent environmental variables were retained after analysis of collinearity (Appendix 4: Fig. S1), and were used to build a global model including: catchment area; morpho-edaphic index (MEI); percentage area of lakes and ponds within the hydrological basin of the catchment; percentage of land in the catchment area covered by coniferous forest (Table 1). The two best models selected evaluated the effect of MEI (Model 1; Table 1) and the additive effect of MEI and catchment area (Model 2; Table 1) on Arctic charr site occupancy. Catchment area had a positive, but not significant impact on Arctic charr occupancy, (Model 2: p = 0.196; Table 1), whereas MEI had a highly significant negative effect in both best models selected (Model 1 and 2: p < 0.000; Table 1; Fig. 5).
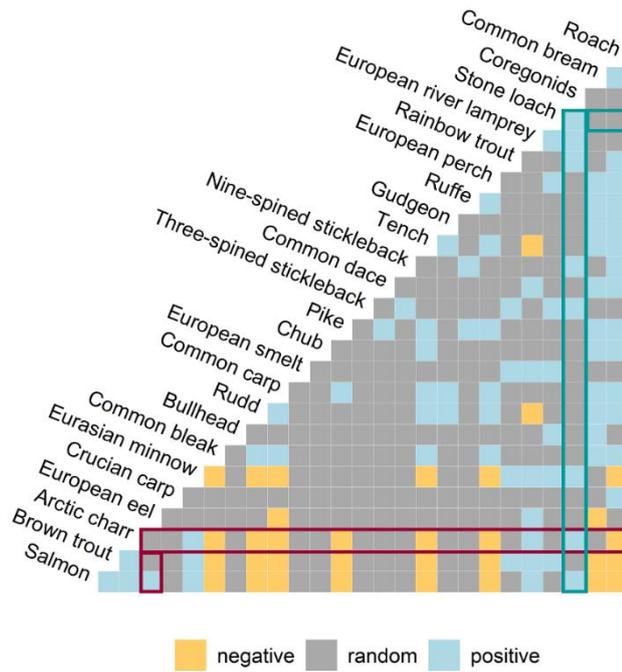
Figure 4. Heat map showing pairwise fish species associations of probabilistic co-occurrence determined by eDNA metabarcoding presence-absence data ($n = 101$ lakes). Significant ($p < 0.05$) associations for Arctic charr and coregonids are highlighted in dark magenta and dark seagreen respectively.

Table 1. Results of GLM analysis showing the effects of environmental variables on Arctic charr site occupancy. The factors were standardised and centred for the full model. The best models (Model1 and Model2) were selected using *quasi*-AIC values. Bold values indicate significant contributions.

Term	Estimate	Standard error	Test statistic	P
Full model				
Catchment area	0.41358	0.30626	1.350	0.180
MEI	**-3.03524**	**0.63111**	**-4.809**	**6.05e-06 *****
Lake area in catchment	-0.09882	0.34915	-0.283	0.778
Pond area in catchment	-0.52505	0.73068	-0.719	0.474
Land covered by coniferous	-0.16223	0.37614	-0.431	0.667
Model1: MEI	**-1.4206**	**0.2831**	**-5.018**	**2.47e-06 *****
Model2: MEI	**-1.382e+00**	**2.867e-01**	**-4.821**	**5.54e-06 *****
Model2: Catchment area	1.161e-05	8.918e-06	1.301	0.196
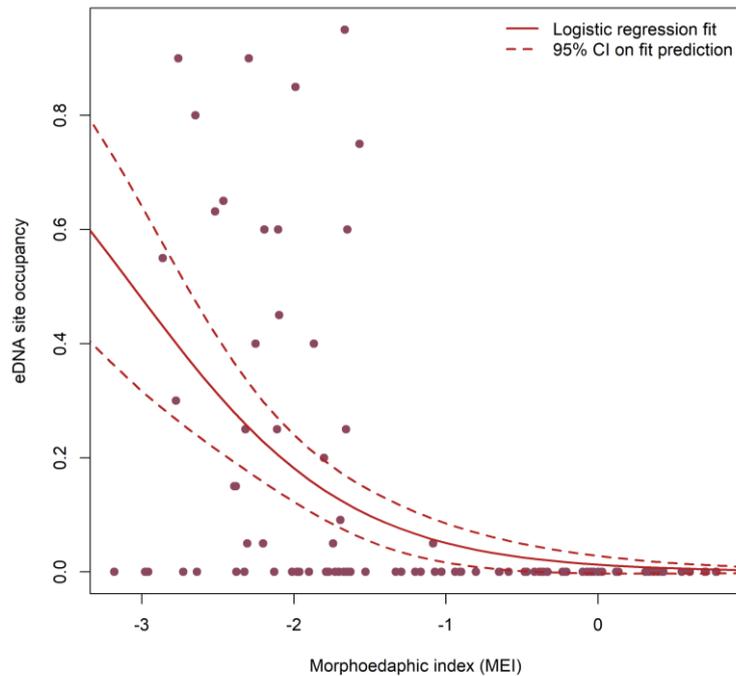
Figure 5. Scatterplot of observed data of Arctic charr eDNA site occupancy and morpho-edaphic index (MEI) values for the lakes surveyed with a fitted prediction line and 95% confidence intervals.

5.5 Discussion

We estimated distribution and abundance of Arctic charr and two coregonid species (European whitefish and vendace) using fish eDNA metabarcoding data and compared these data to historical data from conventional surveys from 101 UK freshwater bodies. We found consensus between eDNA and historical data on Arctic charr and coregonids distribution and, overall, only eight sites out of the 101 surveyed showed differences of presence/absence and ten sites showed moderate or low agreement in terms of abundance estimates. Significant positive correlations were found between abundance estimates from eDNA site occupancy and historical data for Arctic charr and coregonid species. Additionally, based on eDNA fish community detections, a number of significant positive and negative species associations were found, indicating that fish composition influences the distribution of the studied species. The morpho-edaphic index (an index of lake productivity) was identified as the most significant abiotic predictor of Arctic charr abundance in UK lentic systems. Such index could be used, together with other biological parameters (i.e. fish community), to predict the presence of Arctic charr in low productivity lakes.

Our eDNA results have contributed to improve the current knowledge of Arctic charr and coregonids distribution and abundance in UK freshwaters. The present study highlights how the implementation of eDNA-based methods could improve *in-situ* and *ex-situ* management of these threatened fish and help environmental managers to identify habitats of primary importance for the species conservation. In addition, this study showed that eDNA metabarcoding could be broadly applied to assess the abundance and distribution of these three priority conservation fish in UK lentic systems, and be used in conjunction with conventional monitoring methods when other parameters (e.g. age structure or sex ratio) are required for the assessment of lacustrine fish populations.

Distribution of priority conservation fish species based on eDNA observation and historical data

Our results add to a growing body of evidence of the sensitivity of eDNA-based methods for monitoring low-density populations (Sigsgaard *et al.*, 2015; Pfleger *et al.*, 2016; Thomsen *et al.*, 2012a; Janosik and Johnston, 2015; Boothroyd *et al.*, 2016).

Arctic charr is known, through conventional surveys, to occur in 27% of the sites surveyed. Metabarcoding of eDNA detected Arctic charr in all of these sites, and an additional four sites (30% of sites in total) including Loch Lomond, Loch Lubhair, Loch Skeen and Sprinkling Tarn (vendace refuge site). In Loch Lomond, Loch Lubhair and Loch Skeen estimated abundance of Arctic charr from eDNA site occupancy was low, suggesting that either this species is rare within these freshwater bodies or potential false positives. For example, given the low number of genetic sequences assigned (maximum 11 reads after applying the threshold), detections of Arctic charr in Loch Skeen could be false positives arising from barcode misassignment or tag jumps that could have occurred as samples were not sequenced using unique combinations of indexed primers, in line with the library preparation protocol used in Chapter 2 and 3 (Deakin *et al.*, 2014; Schnell, Bohmann and Gilbert, 2015). Furthermore, recent netting surveys did not find Arctic charr in the loch (Winfield *et al.*, 2008b). More interesting is the unexpected eDNA detections of Arctic charr in Loch Lomond and Lubhair, which are likely to represent true records as number of reads assigned to this species were

higher (566 total reads assigned and detections at four sites in Loch Lubhair; 1,303 total reads assigned and detections at three sites in Loch Lomond). Indeed, there are two historical records of Arctic charr in Loch Lomond dating to the late 1800s (Adams, 1994), but this species has not been detected since, whereas there are no historical records of Arctic charr in Loch Lubhair, but the possibility of the species dwelling in this loch has not been excluded (C.W.B. and C.E.A. personal communication). To confirm positive metabarcoding detections when these are inconsistent or dubious (i.e. low read counts or site occupancy) species-specific molecular assays (Harper *et al.*, 2018; Bylemans *et al.*, 2019) or the use of separately sequenced PCR replicates in metabarcoding analysis can be used to address this issue (Ficetola *et al.*, 2015; Alberdi *et al.*, 2018). For instance, we have used an Arctic charr qPCR assay (Chapter 4) to confirm the positive detection of this species in one sample from Kielder Water (Appendix 4: Fig. S2). Alternatively, sampling during the species' spawning season or use a higher number of biological and technical replicates might also increase the probability of detection (see Chapter 2, 3 and 4).

Excluding these few discrepant results, we observed a consensus between expected distribution of Arctic charr and eDNA records. This consistency reflects a substantial reliability of eDNA detections for this low-abundance species. Therefore, these preliminary results from eDNA may help to identify sites that require further investigation, especially when uncertain molecular detections are found in poorly monitored sites. Furthermore, eDNA surveys can also identify sites where endangered species have successfully established new populations (Janosik and Johnston, 2015; Cowart *et al.*, 2018; Hunter *et al.*, 2018) such as our observations of consistent detections of Arctic charr in Sprinkling Tarn. The high eDNA site occupancy in the tarn (positive detections in 80% of the sampling locations) suggested the establishment of an Arctic charr population whose origin is still unknown given that this tarn has not been monitored with conventional surveys since 2005 and this species was previously unrecorded in the tarn (Lyle, Maitland and Winfield, 2006).

Coregonids were detected in Loch Awe and in Windermere, outside their known distribution. In both cases occupancy was very low (one positive sample per site). These detections most likely arise from environmental contamination as coregonids

(e.g. *Coregonus autumnalis*) are commonly used as live-bait by recreational anglers (I.J.W. and C.W.B. pers. comm.). Supporting the hypothesis of environmental contamination is the lack of consistent coregonids detections across different annual eDNA sampling campaigns in Windermere (Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019) or during the periodic monitoring surveys occurring in the lake (Winfield *et al.*, 2007; Winfield *et al.*, 2008a). The presence of DNA in the environment does not necessarily reflect the presence of living organisms, especially when eDNA detections are not consistent (Jerde, 2019). However, previous eDNA surveys demonstrated the ability of this tool to detect unrecorded species at the early stage of colonisation, and only later identified in the environment using conventional surveys (see Chapter 3; Klymus *et al.*, 2017; Blackman *et al.*, 2017; Holman *et al.*, 2019), therefore we cannot exclude *a priori* that the low eDNA detections found in this study could represent true records.

The broad agreement observed in the present study between eDNA metabarcoding results and historical data from traditional surveys should also be evaluated in the light of the considerable different effort for data collection between the two methodologies. The eDNA results presented here were largely collected over a two year period of broad-scale surveys across the UK from a small team of researchers and collaborators, whereas the historical data are based on much longer-term information gathered from conventional methods that are more labour intensive in the field, and potentially damaging to endangered species populations, hence they cannot be applied broadly or periodically. This remarkable difference in sampling effort and resources used empowers the eDNA metabarcoding as a much more cost-effective tool for periodic and broad-scale monitoring surveys of threatened species (Rees *et al.*, 2014; Deiner *et al.*, 2017). Environmental stressors and climate change are currently outpacing the ability of containing the decline of such species, therefore, more frequent surveys would likely prevent these species' extirpations allowing prompt conservation actions (Winfield *et al.*, 2008a; Winfield *et al.*, 2010; Elliott and Bell, 2011; Winfield *et al.*, 2012; Winfield *et al.*, 2013).

Abundance estimates from eDNA metabarcoding and conventional surveys

In line with previous studies (Hering *et al.*, 2018; Li *et al.*, 2018), we demonstrated that eDNA metabarcoding provides results which can be easily adapted to fit some of the current metrics used for fish population monitoring. For example, the Water Framework Directive requires the implementation of estimates of fish abundance to assess the ecological status of freshwater bodies. Li *et al.*, (2018) demonstrated that a five-level DAFOR scale, initially adopted for surveying aquatic macrophytes (Pall and Moser, 2009), can be adapted to estimate fish relative abundance and facilitate the integration into WFD approaches. Here, we have used a similar approach to assess the accuracy of this method and obtain estimates of abundance for Arctic charr and coregonids in a variety of UK lentic systems.

Overall, abundance scores from historical data and eDNA site occupancy showed good agreement, and we observed a strong positive correlation between methods for coregonids, and a weaker, but still significant correlation for Arctic charr. Strong agreement of abundance estimates between methods was found in 63% of the sites surveyed where Arctic charr is known to occur and 45% of the sites within the distribution range of coregonids. Such agreement could be even higher if data from conventional surveys and eDNA had been gathered simultaneously.

Most of the sites surveyed with low or moderate agreement between abundance estimates from eDNA and historical data showed higher abundance scores assigned by historical data (five sites for Arctic charr and three sites for coregonids; Fig. 1; Appendix 4: Table S1). In metabarcoding studies, estimates of species relative abundance can be impacted by inadequate sampling effort (Evans *et al.*, 2017; Grey *et al.*, 2018; Cantera *et al.*, 2019), eDNA degradation (Lacoursière-Roussel *et al.*, 2016; Tsuji *et al.*, 2017; Jo *et al.*, 2019) or low level of in-laboratory technical replication and sequencing depth (Smith and Peay, 2014; Ficetola *et al.*, 2015). Hence, eDNA metabarcoding protocols, similarly to other biomonitoring methods, need to be adapted to the bio-geographical regions and target species surveyed. The standardised eDNA metabarcoding protocol used across the sites surveyed and the general agreement found between methods leads to the hypothesis that high abundance

scores assigned from historical data might be biased due to a lack of monitoring surveys in recent years in populations that might be declining. Conventional surveys are costly in terms of time, financial resources, manpower skills and equipment and, therefore, not frequently manageable. Additionally, their invasive or destructive nature limits their repeated applicability. Thus, the integration of eDNA methods into periodic long-term monitoring surveys appears essential to identify areas requiring management actions, especially when low agreement between assessment methods occur, and in sites of conservation importance (e.g. SSSIs). For example, Loch Doon and Llyn Padarn are sites of conservation interest for Arctic charr, and the species estimated abundance from eDNA was lower than expected from historical data. The Arctic charr population in Loch Doon is threatened by acidification and a strong concern exists on Arctic charr population in Llyn Padarn, which was threatened by eutrophication (Maitland *et al.*, 2007; Maitland and Lyle 2013). The low abundance estimates assessed by eDNA metabarcoding might indicate that Arctic charr populations at these sites of conservation interest are at risk, thus demanding further implementation of conservation plans. Similarly, low abundance scores were estimated by eDNA metabarcoding compared to expected abundance estimated from historical data for vendace in Derwent Water and whitefish in Brothers Water. Given that the population of whitefish in Brothers Water is considered to be in unfavourable condition (Winfield *et al.*, 2011), and Derwent Water represents the last consistently vendace recruiting site in the UK (Winfield *et al.*, 2004), further monitoring is required to assess the status of these two populations and ensure that they are not exposed to additional threats, such as the further introduction on non-native species.

Evaluating translocation plans and factors influencing Arctic charr and coregonids distribution and abundance

Habitat restoration is generally the favoured approach to safeguard endangered populations and re-establish functional systems (Winfield *et al.*, 2012; Maitland and Lyle, 2013). For example, to mitigate eutrophication threatening Arctic charr populations, restoration has been carried out in Windermere (Winfield *et al.*, 2008a). The implementation of restoration measures, when possible, is often costly and restored habitats may require many years to fully recover. The timescales required for

recovery may mean that the decline of native, sensitive populations may continue to the point where they are lost. To preserve fish populations from extinction, the translocation of fish to new refuge sites is often the only option available (Seddon *et al.*, 2014). However, translocated fish often fail to establish new self-sustaining populations and, because post-release monitoring is rarely performed, and the time and cause of failure often remain unknown (Beck *et al.*, 1994; Seddon *et al.*, 2007). We provide further evidence that translocations of Arctic charr in Kielder Water and Grimwith Reservoir, as well as vendace released in Sprinkling Tarn, have failed to establish new sustaining populations at these refuge sites. In Kielder Water, Arctic charr DNA was detected in one water sample of 11 collected (Appendix 4: Table S1), and this result was confirmed using targeted qPCR (Appendix 4: Fig. S2). Given the low number of positive detections and the relatively recent translocation of Arctic charr adults to this reservoir, this detection might indicate that some individuals survived within the water body, but failed to establish a self-sustaining population.

This study also confirmed that translocated vendace in Sprinkling Tarn failed to establish, and found consistent detections of the previously unrecorded Arctic charr (see Lyle, Maitland and Winfield, 2006; Appendix 4: Table S1). The abundant Arctic charr population in the tarn might be one of the reasons explaining the failed survival of vendace at this refuge site. Coregonids and Arctic charr are known to co-occur in sites where they can occupy different niches or alter their foraging behaviour, otherwise the use of similar resources and habitats can lead to competitive exclusion (Sandlund *et al.*, 2009; Eloranta *et al.*, 2011). Our observations confirm that eDNA metabarcoding is a valuable tool for evaluating the establishment of new populations at their refuge sites, and can be conveniently deployed in difficult-to-access sites for post-release and long-term monitoring of translocation programmes.

This whole-community, non-invasive genetic tool can evaluate causes of a failed establishment and, most importantly it can be used to assess the community composition of water bodies prior to fish translocations as a number of factors has to be evaluated when choosing appropriate refuge sites including abiotic and biotic conditions (i.e. species composition). For example, we observed negative and positive associations between Arctic charr, coregonids and other fish species (Fig. 4). These

interactions can be used to select suitable sites for future translocations or predict variation of distribution of Arctic charr and coregonids based on changes of fish species composition. For example, roach has been repeatedly implicated in declines of native Arctic charr populations (Winfield and Durie, 2004; Winfield *et al.*, 2008). Roach is often used as a live-bait and this practice can determine the release of living specimens in sites where this species does not naturally occur determining, when conditions are favourable, the establishments of new populations (Winfield *et al.*, 2008). Expansion of introduced roach populations have been documented and are considered a threat to native Arctic charr and coregonid populations in lentic systems (Winfield *et al.*, 2008a). The sensitivity of eDNA methods for early detection of introduced non-native species (Jerde *et al.*, 2011; 2013; Mahon et al, 2013; Takahara *et al.*, 2013), could be used to monitor newly introduced competitors and their spread in waterbodies that support native Arctic charr and coregonids. Determining the distribution and abundance of roach and that of other potential competitors or predators of Arctic charr and coregonids (Fig. 4) is as important as understanding distribution and abundance of these priority conservation species. For example, when a new translocation programme is planned, fish community data from eDNA metabarcoding can be used to complement ecological and physical data, and provide crucial information on the most suitable candidate sites for translocation.

We additionally observed that the environmental factor MEI (i.e. morpho-edaphic index) is a significant predictor of Arctic charr abundance and it is negatively associated to Arctic charr site occupancy at the sites surveyed in this study (Table 1, Fig. 5). MEI is effectively a standard proxy for total phosphorus concentrations in UK lakes, and is used to monitor the ecological status of macrophytes following standard metrics of assessment in line with the WFD (UKTAG, 2006; Environment Agency, 2012). High values of MEI indicate good ecological status of lake macrophytes and MEI can be therefore considered as a productivity index where high values are typically associated to eutrophic lakes. MEI is calculated as the lakes' median alkalinity over the lakes' mean depth, and given Arctic charr preferences for deep and oligotrophic lakes we can explain the ecological value of MEI as a good predictor of Arctic charr abundance in UK lentic systems. High values of MEI associated with shallower and more eutrophic lakes will represent less suitable habitats for Arctic charr populations and, therefore, this

index might be integrated into future evaluation of appropriate habitats for Arctic charr translocations or for the assessment of sites where this species naturally occurs.

5.6 Conclusion

Environmental DNA metabarcoding data are scalable and adaptable to current metrics for monitoring fish communities (Hering *et al.*, 2018; Li *et al.*, 2019a) and, despite some limitations (i.e. interpretation of inconsistent detections of unrecorded species), these results are generally accurate for detections and semi-quantitative estimates of low-abundant species. Overall, our fish eDNA metabarcoding results are also in line with the expected distribution and abundance of Arctic charr and coregonids in UK lentic systems. This non-invasive, whole-community tool additionally provides insights on biotic and abiotic factors influencing the actual and future distribution and abundance of these threatened species. Thus, the use of eDNA metabarcoding would benefit these species management through an accurate assessment of native and translocated populations. The use of species translocations as a management tool to combat the impacts of climate change and other anthropogenic stressors which threaten native populations is likely to increase in future years. We suggest that eDNA metabarcoding can assist with the choice of future refuge sites for species of conservation interest, reduce the costs of post-release monitoring and, simultaneously, provide crucial information on difficult-to-sample sites, hence representing a valuable tool to guide and monitor outcomes of conservation actions.

Chapter 6

General discussion

6.1 Environmental DNA metabarcoding provides reliable estimates of fish biomass and abundance

Resources for monitoring threatened fish species are generally limited so obtaining the most reliable estimates of species abundance from cost-efficient, non-invasive sampling methods is essential to provide accurate management guidelines and determine conservation priorities (Maxwell and Jennings, 2005). The action of sampling involves obtaining a group of specimens or objects from a larger population of which they are representative, and each fish sampling method has its own biases in providing quantitative information of fish biomass/abundance (Deagle *et al.*, 2019). Therefore, what should be evaluated is the methodology's ability to accurately represent the biomass/abundance of the population from which the samples were obtained (Hering *et al.*, 2018). Chapter 2 empirically demonstrated the existence of strong and significant associations between eDNA metabarcoding data (read counts and site occupancy) and absolute fish biomass/abundance. To date, assessments of the accuracy of quantitative estimates from eDNA metabarcoding in natural systems, have resulted from comparisons between this molecular approach and indirect estimates of relative abundance from conventional methods (Hänfling *et al.*, 2016; Li *et al.*, 2019a; Pont *et al.*, 2018; Sard *et al.*, 2019). However, this thesis demonstrated that eDNA metabarcoding read counts and site occupancy reflect the absolute fish abundance and biomass obtained from direct measures of count and weight data for the whole fish community in the fishery ponds used as a case study. In large and heterogeneous lentic systems, environmental factors (i.e. hydrological conditions) and fish ecology and life history (i.e. spawning) might influence real abundance patterns, therefore eDNA sampling campaigns should be planned accounting for these variables in order to achieve accurate and reliable estimates of abundance (Chapter 3; de Souza *et al.*, 2016; Eiler *et al.*, 2018; Lawson Handley *et al.,* 2019). For example, eDNA sampling in lentic systems should be carried out when water is well mixed and taking into account the species' breeding period, as stratification or spawning activity might

misrepresent the real species' abundance patterns (Chapter 3; Hänfling *et al.*, 2016; Lawson Handley *et al.*, 2019).

Policy makers, stakeholders and environmental managers require evidence of the capability of eDNA metabarcoding protocols to provide consistent, sensitive and repeatable quantitative estimates for the implementation of this tool into current fish monitoring surveys (Collen *et al.*, 2009). This thesis has shown the importance of the collection of multiple eDNA samples and filter replicates to increase the sensitivity of eDNA metabarcoding and enhance species detection probabilities and accuracy of abundance estimates (Chapter 2). The consistency observed between qualitative and quantitative outcomes of different eDNA filtration (0.45 μm Sterivex filters vs. 0.45 μm MCE filters) and preservation strategies for filters (buffers vs. freezing) demonstrated that the eDNA metabarcoding protocol used in fish is also robust and repeatable, at least for the filter's pore size and materials evaluated within this work (Chapter 2). The possibility of using different processing strategies for eDNA-based studies and obtaining consistent estimates of fish biomass/abundance provides a great opportunity for managers to guarantee repeatable qualitative and quantitative estimates in the absence of a standard collection/filtration approach.

Environmental managers would also benefit from using indicators of abundance trends that are comparable on large-scale and sustainable over time, accounting for available resources and the potential impacts on the surveyed populations. The ease of eDNA sampling and its non-invasive approach make this tool broadly accepted and adequate for landscape scale monitoring of threatened species (Chapter 5; Rees *et al.*, 2014; Mize *et al.*, 2019; Sales *et al.*, 2019). In addition, when monitoring abundance at landscape scale, abundance estimates should be comparable across a variety of habitats, even when uneven sampling efforts are employed (Hering *et al.*, 2018). This thesis has demonstrated the accuracy of abundance estimates from eDNA metabarcoding read counts, however, these abundance estimates would require the use of correction factors for intercalibration across studies, such as the inclusion of internal standards of known copy number during sequencing (e.g. Ushio *et al.*, 2018). Conversely, the proportion of area occupied by fish species, namely the site occupancy, is easily scalable to fit standardised metrics of abundance (Chapter 5;

Hering *et al.*, 2018; Li *et al.*, 2019a). This thesis has therefore demonstrated that abundance estimates from fish eDNA metabarcoding site occupancy are adequate for long-term or broad-scale monitoring programmes, such as the Water Framework Directory (Hering *et al.*, 2018; Li *et al.*, 2019a) and that these scalable abundance estimates are also accurate for the assessment of the status of multiple threatened fish populations across a variety of UK lentic systems (Chapter 5).

Current fish monitoring methods (i.e. fyke-nets, gill-nets or electrofishing) often inaccurately represent fish population size, and the efficacy of these methods often varies with target species (Hardie *et al.*, 2006; Basler and Schramm, 2011; Lintermans, 2016). Moreover, conventional fish sampling methods cannot be extensively deployed to cover the entire area of interest (either between or within freshwater habitats), thus estimates of fish biomass/abundance are likely less accurate especially towards low-abundant species (Winfield *et al.*, 2007; Dennerline *et al.*, 2012). The sensitivity and manageability of eDNA metabarcoding approaches for long-term monitoring of fish communities is not only essential for monitoring threatened or less abundant species, but also to predict or assess loss or changes in population size of these sensitive taxa resulting from early detections of non-native competitive species (Ficetola *et al.*, 2008; Jerde *et al.*, 2011; Dejean *et al.*, 2012; Goldberg *et al.*, 2013). For example, the expansion of roach (*Rutilus rutilus*) and bream (*Abramis brama*) populations in Windermere is already contributing to the decline of native Arctic charr populations (Chapter 3; Winfield *et al.*, 2008a; Hänfling *et al.*, 2016). The identification of ruffe (*Gymnocephalus cernua*) in Windermere is also likely to impact the threatened Arctic charr within the lake, but the early detection of this competitive species from eDNA metabarcoding could be used to put conservation plans in place and help slow down or contain the negative impact of this introduced species on native fish populations in Windermere (Chapter 3; Hänfling *et al.*, 2016).

6.2 Environmental DNA metabarcoding determines spawning location and timing of endangered fish species

Measuring trends in abundance is just one of several important indicators in the assessment of fish populations (Radinger *et al.*, 2019). Species' biology information

such as spawning timing, location and output are crucial considerations for effective fisheries management (Koenig *et al.*, 2000; King *et al.*, 2010; Erisman *et al.*, 2017). Fish species are highly mobile and selective in their habitat preferences and, even within confined lacustrine systems, some species undergo migration to aggregate in specific locations and at discrete times to breed. Thus, the protection of such critical areas is broadly recognised as a focal point for conservation efforts (Winfield, 2004; Winfield *et al.*, 2015; Erisman *et al.*, 2017). Arctic charr and coregonids (i.e. vendace and European whitefish) are lithophilic species that require hard substrates for spawning. As a result of these preferences, Artic charr and coregonids will likely only use a fraction of the total available habitat when spawning (Low *et al.*, 2011; Winfield *et al.*, 2015). These deep-water species usually occupy the offshore areas of lacustrine systems and migrate to shallow shore waters to spawn (Miller *et al.*, 2015; Lyle *et al.*, 2019). Arctic charr and coregonids additionally show regional differentiation, often genetically-based, which leads to characteristic spawning adaptations that require local monitoring to ensure adequate conservation through the preservation of successful recruitment for each discrete population (Garduño-Paz *et al.*, 2010; Corrigan *et al.*, 2011). This thesis used the autumn-spawning population of Arctic charr from Windermere (England, UK) as a case study to demonstrate that read counts from eDNA metabarcoding reflect locations and the timing of spawning activity (Chapter 3). Coregonids and Arctic charr are biologically similar, thus the fish eDNA metabarcoding protocol used within this thesis might also be advantageous to monitor spawning activity of coregonid species in freshwater habitats. Identification of the accurate location of coregonid spawning grounds would greatly contribute to the management of these threatened fish species and could be used to create selected areas of special protection within lacustrine systems and to safeguard the critical spawning event. The protection of limited areas (i.e. fish spawning sites) within large freshwater environments would ensure the preservation of habitats' characteristics that are essential to support threatened species spawning activity. For example, previous studies have shown that siltation or sedimentation determine a deterioration of the already limited Arctic charr spawning grounds in lakes (Igoe *et al.*, 2003; Winfield *et al.*, 2009; Miller *et al.*, 2015). Successful management of these limited areas would save resources as conservation efforts could be limited to specific focal sites. The protection

of fish spawning areas from human impact is a common strategy of fisheries management in marine habitats (Erisman *et al.*, 2017), and a similar approach would benefit the conservation of threatened freshwater fish (Adams *et al.*, 2006; Corrigan *et al.*, 2011).

In addition to the identification of spawning areas, long-term eDNA monitoring of species spawning might help us to better understand interannual variation of spawning effort and recruitment, although the act of spawning does not necessarily translate into an efficient recruitment (MacKenzie *et al.*, 1998; Marshall *et al.*, 1998). Future studies could evaluate the ability of eDNA metabarcoding to assess fish species recruitment by for example, monitoring species-specific variation of eDNA at the spawning grounds after hatching. The hatching of both Arctic charr and coregonids is expected around March/April and post-hatching, young fish use the littoral habitats close to their native grounds where they feed among the stones and gravel, which also provide shelter against predation (Klemetsen *et al.*, 2003). Species-specific peaks of eDNA in the coarse-substrate native habitats after hatching would likely be associated with abundance of young individuals at those sites thus providing crucial information regarding monitoring spawning activity and recruitment of these priority conservation fish.

6.3 Environmental DNA metabarcoding effectively monitors UK priority conservation fish species in sensitive sites

This thesis demonstrated that eDNA metabarcoding can be as sensitive as targeted molecular assays for detection of low-abundant species (e.g. rare or threatened species) when using equal number of PCR replicates (Chapter 4). The high number of technical replicates used (*n*=12) for maximising detection of low-density species (Chapter 4) would be expensive for large-scale monitoring, however, this work has shown that target species detection probability is influenced by the original concentration of DNA in water samples and additionally, that these concentrations can be used to evaluate the number of technical replicates needed based on the type of samples that are to be processed (e.g. shore *vs.* offshore samples). For example, the collection of shoreline samples is sufficient for detection of Arctic charr and coregonids during the species' spawning period and, these samples with usually higher

concentration of DNA, would require a lower number of PCR replicates (Chapter 4). Conversely, the collection of offshore samples might be necessary to enhance detection probabilities of Arctic charr and coregonids when sampling outside the spawning seasons (Chapter 5; Lawson Handley *et al.*, 2019), and a higher number of PCR replicates might be necessary to process offshore samples with lower concentrations of original DNA template and to increase detection probability of the target species.

The fish eDNA metabarcoding protocol used within this thesis was accurate to provide relative abundance data of UK priority conservation fish in sites of special interest (i.e. Special Area of Conservation or refuge sites; Chapter 5). Environmental DNA metabarcoding results are a reliable and sensitive tool to guide conservation efforts and species management in a number of situations such as eradication programmes, assessment of suitability of refuge sites through species composition analysis, post-translocation evaluation of species establishment, long-term monitoring of established self-sustaining populations in refuge sites, and recovery of fish communities after habitat restoration. Furthermore, fish community composition and site-specific factors can be used together with relative abundance estimates from eDNA to evaluate present and future distributions of threatened species (Chapter 5) whereby site occupancy models could be used to predict local extinction in poor quality habitats or colonisation of new sites with suitable characteristics for establishment of new populations (MacKenzie and Nichols, 2004).

6.4 Closing remarks: understanding limitations and future directions of eDNA

This thesis showed that eDNA approaches can significantly improve the capacity to monitor priority conservation fish in UK lentic systems, however, despite the promising applications of eDNA in ecology and conservation, this tool should not be considered as a stand-alone method (Beng and Corlett, 2020). The sampling of eDNA cannot replace conventional surveys because both types of monitoring methods deliver different information and they should, therefore, complement each other. For example, organism capture allows the collection of additional and important non-

genetic data, such as size, age, maturity, isotopic measurements, enzymatic activity, and hormone activity (Lawson Handley, 2015). Researchers, managers and stakeholders should consider which method is the most appropriate based on the questions they want to address (Qu and Stewart, 2019). The recognition of the limitations of eDNA approaches allows a critical interpretation of the results and to draw appropriate conclusions (Cristescu and Hebert, 2018). As shown in this thesis, one of the main challenges of eDNA data is the interpretation of inconsistent detections of unrecorded species, providing that it can be difficult in such cases to distinguish between the detection of DNA and the detection of a living organism within the environment (Jerde, 2019; Beng and Corlett, 2020). The current lack of a clear understanding about the ecology and the dynamics of eDNA (origin, state, transport and fate) can limit the inferences on the organism's presence when molecular detections are not robust across temporal, spatial or technical replicates (Deiner and Altermatt, 2014; Barnes and Turner, 2016; Collins *et al.*, 2018; Seymour *et al.*, 2018). The eDNA workflow (sampling, DNA preservation, laboratory and analysis protocols) also influences the accuracy of the results (Coissac, Riaz and Puillandre, 2012; Eichmiller *et al.*, 2016; Goldberg *et al.*, 2016). Some of these limitations can be overcome by optimising the experimental workflows based on the target taxa and their ecology, study sites and environmental conditions (Cristescu and Hebert, 2018). Moreover, in metabarcoding studies, the correct identification of genetic sequences to low taxonomic levels is entirely dependent on the genetic information included in the reference databases (Lecaudey *et al.*, 2019; Doble *et al.*, 2020; Schenekar *et al.*, 2020). The incompleteness of public reference databases such as GenBank nucleotide database at the National Center for Biotechnology Information (NCBI), the European Molecular Biology Laboratory (EMBL) database or the Barcode of Life Database (BOLD), often impedes the assignment of a large part of the genetic sequences in eDNA samples (Collins *et al.*, 2019; Sales *et al.*, 2019; Apothéloz-Perret-Gentil *et al.*, 2020). To circumvent this issue some studies rely on custom reference database including barcoded sequences of the local fauna of interest, such as the 12S custom fish database used in this thesis and developed in Hanfling *et al.*, 2016 (but also Cilleros *et al.*, 2019; Shen *et al.*, 2019). Even in such cases, the use of a single marker might not be enough for the accurate description of all taxa within a group at the species level. For

example, in this thesis the 12S region used for metabarcoding analysis cannot distinguish between two species of the Percidae family (*Perca fluviatilis* and *Sander lucioperca*) nor between two species of the *Coregonus* genus (*C. lavaretus* and *C. autumnalis*) (see Chapter 3). When there are no accurate background information of the sites surveyed and the research interest is the correct identification of genetic sequences to the species level, the use of more than one marker can be a solution (Giebner *et al*., 2019; Morey *et al*., 2020; Schenekar *et al*., 2020). However, this practice will certainly increase the cost and the effort of eDNA metabarcoding surveys and, to date, the reduced cost and the increased speed remain the key advantages of this whole-community tool (Sigsgaard *et al.*, 2015).

Beyond the limitations of eDNA metabarcoding, this thesis has demonstrated that this tool can effectively complement conventional monitoring methods for example providing accurate quantitative information of fish communities including low-abundant species, monitoring the species spawning activity and assessing priority conservation fish at their native and refuge sites or in sites of conservation interest. Furthermore, this thesis has shown the potential of using the information provided by eDNA metabarcoding to evaluate present and future distribution and abundance of threatened fish in UK lentic environments. Overall, this work should instill more confidence in managers to use eDNA metabarcoding, along with other monitoring methods, for future fish assessments and to plan conservation actions.

The whole-community approach used here, cannot represent an endpoint and, this rapidly developing tool for biodiversity monitoring should be a key focus of future research, especially investigations that consider the use of eDNA in population genetics. This concept has recently emerged in the field of eDNA with a handful of outstanding studies demonstrating the ability to obtain intraspecific genetic diversity information from a collection of environmental samples. For example, haplotype variation using mitochondrial DNA has been determined for a number of target species (Sigsgaard *et al.*, 2016; Baker *et al.*, 2018; Parson *et al.*, 2018) as well as in multispecies assays (Elbrecht *et al.*, 2018). The use of eDNA metabarcoding for population genetics inferences requires the use of broad-range primers with enough resolution for haplotype identification, which are also able to capture the largest species diversity

within a certain taxonomic group. The 12S primers used within this thesis (Riaz *et al.*, 2011; Kelly *et al.*, 2014) would not be appropriate to resolve intraspecific variation given the short length of fragment amplified (~106 bp). The amplification of suitable longer regions using broad-range primers is a key area for future research considering fish communities in UK freshwaters as it represents a potential tool for assessing conservation of Arctic charr and coregonids populations in the UK. For example, Arctic charr displays great plasticity worldwide as well as in the UK with a number of, sometimes sympatric, morphotypes showing discrete migration patterns, spawning-time and ecology (Adams and Maitland, 2007; Corrigan *et al.*, 2011). These morphs are often associated to different, yet understudied genotypes that create a mosaic of genetic diversity with fundamental consequences and challenges for conservation (Adams *et al.*, 2003; Adams *et al.*, 2006). In this context, future eDNA studies should also explore the use of nuclear markers or long-read sequencing to reach a greater genomic coverage as this would provide crucial information on the evolutionary processes that have driven this high diversification and this knowledge would, as a consequence, benefit the implementation of appropriate conservation measures (Adams *et al.*, 2019).

The ever-advancing technology is also increasingly moving towards *in-situ* species monitoring with autonomous sampling and analysis devices that can be also operated remotely (Mcquilllan and Robidart, 2017). Such field instruments include completely automatised systems for the collection of eDNA samples and detection of target taxa, from microbes to vertebrates (Preston *et al*., 2011; Yamahara *et al*., 2019; Hansen *et al*., 2020) as well as portable sequencing instruments (Russel *et al.*, 2018). Indeed, the third-generation sequencing offers hand-held systems (i.e. MinIon and SmidgION) with good read length and adaptable to field usage in remote settings (Russel *et al.*, 2018). Although still limited in terms of throughputs, these technologies hold interesting perspectives for future on-site biomonitoring surveys (Goodwin *et al*., 2016). In field surveys of target species might be facilitated by the application of PCR-free molecular diagnostic technology that enhances specificity for target species. For example, the use of RPA-CRISPR-Cas12 has been successful in the detection of *Salmo salar* in eDNA samples (Williams *et al.*, 2019). Community data from shotgun sequencing, not relying on PCR-mediated amplification steps, can also help overcoming the biases associated

with the use of primers and PCR (Pedersen *et al.*, 2016; Parducci *et al.*, 2019), although the cost for acquiring data with enough depth are still prohibitive, as the technology advances and cost goes down this approach might become a common biomonitoring method (Ruppert *et al.*, 2019).

The simplicity of the sampling of eDNA has also opened new avenues for citizen science programmes (Biggs *et al.*, 2015) using the engagement of citizen in the collection of eDNA samples along with the use of commercial sampling kits from companies specialised in eDNA analysis (e.g. NatureMetrics, Spygen, GENIDAQS). This public engagement into biodiversity monitoring empowers the educational use of eDNA and increases awareness about biodiversity and its conservation (Deiner *et al.*, 2017).

Overcoming at least some of the limitations of eDNA methods in a foreseeable future will be possible with an investment in technology, expansion of reference database and an enhanced understanding of the eDNA dynamics. As we move towards a next-generation monitoring era, it is predictable that eDNA will become a standard tool for biomonitoring, and coupled with conventional methods, will accelerate the understanding of biodiversity changes allowing prompt conservation actions and management decisions to be taken in order to mitigate the effects of global climate change and other anthropogenic pressures.

References

Adams, C.E., 1994. The fish community of Loch Lomond, Scotland: its history and rapidly changing status. Hydrobiologia, 290(1-3), pp.91-102. https://doi.org/10.1007/BF00008956

Adams, C., Fraser, D., McCarthy, I., Shields, S., Waldron, S. and Alexander, G., 2003. Stable isotope analysis reveals ecological segregation in a bimodal size polymorphism in Arctic charr from Loch Tay, Scotland. Journal of Fish Biology, 62(2), pp.474-481. https://doi.org/10.1046/j.1095-8649.2003.00044.x

Adams, C.E., Hamilton, D.J., Mccarthy, I., Wilson, A.J., Grant, A., Alexander, G., Waldron, S., Snorasson, S.S., Ferguson, M.M. and Skúlason, S., 2006. Does breeding site fidelity drive phenotypic and genetic sub-structuring of a population of Arctic charr?. Evolutionary Ecology, 20(1), pp.11-26. https://doi.org/10.1007/s10682-005-2489-4

Adams, C. E., Bean, C. W., Fraser, D., Maitland, P. S., 2007. Conservation and management of the Arctic charr: a forward view. Ecology of Freshwater Fish, 16(1), pp.2-5. https://doi.org/10.1111/j.1600-0633.2006.00180.x

Adams, C. E., and Maitland, P. S., 2007. Arctic charr in Britain and Ireland–15 species or one?. Ecology of Freshwater Fish, 16(1), pp.20-28. https://doi.org/10.1111/j.1600-0633.2006.00179.x

Adams, C.E., Fraser, D., Wilson, A.J., Alexander, G., Ferguson, M.M. and Skulason, S., 2007. Patterns of phenotypic and genetic variability show hidden diversity in Scottish Arctic charr. Ecology of Freshwater Fish, 16(1), pp.78-86. https://doi.org/10.1111/j.1600-0633.2006.00182.x

Adams, C.E. and Maitland, P.S., 2007. Arctic charr in Britain and Ireland–15 species or one?. Ecology of Freshwater Fish, 16(1), pp.20-28. https://doi.org/10.1111/j.1600-0633.2006.00179.x

Adams, C.E., Lyle, A.A., Dodd, J.A., Bean, C.W., Winfield, I.J., Gowans, A.R., Stephen, A. and Maitland, P.S., 2014. Translocation as a conservation tool: case studies from rare freshwater fishes in Scotland. Glasgow Naturalist, 26(Part 1), pp.17-24.

Adams, C.E. and Maitland, P.S., 2018. Arctic Charr in the Lochs of Scotland. Upfront Publishing. 382pp.

Adams, C.I., Knapp, M., Gemmell, N.J., Jeunen, G.J., Bunce, M., Lamare, M.D. and Taylor, H.R., 2019. Beyond biodiversity: can environmental DNA (eDNA) cut it as a population genetics tool?. Genes, 10(3), p.192.
https://doi.org/10.3390/genes10030192

Alekseyev, S. S., Gordeeva, N. V., Samusenok, V. P., Matveev, A. N., Andreev, R. S., Yur'ev, A. L., Smirina, E. M., 2013. Extant and extinct forms of Arctic charr *Salvelinus alpinus* (L.) complex from the Leprindo lake system (Transbaikalia): differentiation in life history, morphology, and genetics. Journal of ichthyology, 53(10), pp.792-803. DOI: 10.1134/S0032945213100020

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

Allard, L., Grenouillet, G., Khazraie, K., Tudesque, L., Vigouroux, R., Brosse, S., 2014. Electrofishing efficiency in low conductivity neotropical streams: towards a non-destructive fish sampling method. Fisheries Management and Ecology, 21(3), pp.234-243. https://doi.org/10.1111/fme.12071

Antognazza, C.M., Britton, J.R., Potter, C., Franklin, E., Hardouin, E.A., Gutmann Roberts, C., Aprahamian, M. and Andreou, D., 2019. Environmental DNA as a non-invasive sampling tool to detect the spawning distribution of European anadromous shads (*Alosa* spp.). Aquatic Conservation: Marine and Freshwater Ecosystems, 29(1), pp.148-152. https://doi.org/10.1002/aqc.3010.

Apothéloz-Perret-Gentil, L., Bouchez, A., Cordier, T., Cordonier, A., Guéguen, J., Rimet, F., Vasselon, V., Pawlowski, J., 2020. Monitoring the ecological status of rivers with

diatom eDNA metabarcoding: A comparison of taxonomic markers and analytical approaches for the inference of a molecular diatom index. Molecular Ecology. https://doi.org/10.1111/mec.15646

Baker, C.S., Steel, D., Nieukirk, S. and Klinck, H., 2018. Environmental DNA (eDNA) from the wake of the whales: droplet digital PCR for detection and species identification. Frontiers in Marine Science, 5, p.133. https://doi.org/10.3389/fmars.2018.00133

Baldigo, B.P., Sporn, L.A., George, S.D. and Ball, J.A., 2017. Efficacy of environmental DNA to detect and quantify brook trout populations in headwater streams of the Adirondack Mountains, New York. Transactions of the American Fisheries Society, 146(1), pp.99-111. https://doi.org/10.1080/00028487.2016.1243578

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L. and Lodge, D.M., 2014. Environmental conditions influence eDNA persistence in aquatic systems. Environmental science & technology, 48(3), pp.1819-1827. DOI: 10.1021/es404734p

Barnes, M.A. and Turner, C.R., 2016. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics, 17(1), pp.1-17.

Barton, K., Barton, M.K., 2020. Package 'MuMIn'in R Version 1.43. 17.

Basler, M.C. and Schramm Jr, H.L., 2006. Evaluation of electrofishing and fyke netting for collecting black carp in small ponds. Transactions of the American Fisheries Society, 135(2), pp.277-280. https://doi.org/10.1577/T05-119.1

Bayer, S.R., Countway, P.D. and Wahle, R.A., 2019. Developing an eDNA toolkit to quantify broadcast spawning events of the sea scallop *Placopecten magellanicus*: moving beyond fertilization assays. Marine Ecology Progress Series, 621, pp.127-141. https://doi.org/10.3354/meps12991

Bean, C.W., 2003a. A standardised survey and monitoring protocol for the assessment of Arctic charr, *Salvelinus alpinus* L., populations in the UK.

Bean, C.W., 2003b. A standardised survey and monitoring protocol for the assessment of whitefish, *Coregonus albula* (L.) and *C. lavaretus* (L.), populations in the UK.

Bean, C., Hall, R. and Thomas, R., 2015. Common standards protocol for population monitoring of Arctic charr, whitefish and vendace. In: JNCC (2015) Common Standards Monitoring Guidance for Freshwater Fauna, Version October 2015, JNCC, Freshwater, ISSN 1743-8160. pp. 33-43

Beck, B.B., Rapaport, L.G., Price, M.S. and Wilson, A.C., 1994. Reintroduction of captive-born animals. In Creative conservation (pp. 265-286). Springer, Dordrecht. https://doi.org/10.1007/978-94-011-0721-1_13

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

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

Bivand, R., Keitt, T., Rowlingson, B., 2019. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-8. https://CRAN.R-project.org/package=rgdal

Blackman, R.C., Constable, D., Hahn, C., Sheard, A.M., Durkota, J., Hänfling, B. and Handley, L.L., 2017. Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples—first record of *Gammarus fossarum* in the UK. Aquatic Invasions, 12(2). https://doi.org/10.3391/ai.2017.12.2.06

Bolgan, M., O'Brien, J., Chorazyczewska, E., Winfield, I.J., McCullough, P. and Gammell, M., 2018. The soundscape of Arctic Charr spawning grounds in lotic and lentic environments: can passive acoustic monitoring be used to detect spawning activities?. Bioacoustics, 27(1), pp.57-85. https://doi.org/10.1080/09524622.2017.1286262

Boothroyd, M., Mandrak, N.E., Fox, M. and Wilson, C.C., 2016. Environmental DNA (eDNA) detection and habitat occupancy of threatened spotted gar (*Lepisosteus oculatus*). Aquatic conservation

Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., Garcia-Vazquez, E., 2017. DNA in a bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. PloS one, 12(9), e0183347.

Bracken, F.S., Rooney, S.M., Kelly-Quinn, M., King, J.J. and Carlsson, J., 2019. Identifying spawning sites and other critical habitat in lotic systems using eDNA "snapshots": A case study using the sea lamprey *Petromyzon marinus* L. Ecology and Evolution, 9(1), pp.553-567. https://doi.org/10.1002/ece3.4777

Brattli, M. B., Egeland, T. B., Nordeide, J. T., Folstad, I., 2018. Spawning behavior of Arctic charr (*Salvelinus alpinus*): Spawning synchrony, vibrational communication, and mate guarding. Ecology and evolution, 8(16), pp.8076-8087. https://doi.org/10.1002/ece3.4277

BRIG, 2007. Report on the Species and Habitat Review (Report by the Biodiversity Reporting and Information Group (BRIG) to the UK Standing Committee). JNCC, Peterborough.

Britton, J.R., Brazier, M., Davies, G.D. and Chare, S.I., 2008. Case studies on eradicating the Asiatic cyprinid *Pseudorasbora parva* from fishing lakes in England to prevent their riverine dispersal. Aquatic Conservation: Marine and Freshwater Ecosystems, 18(6), pp.867-876. https://doi.org/10.1002/aqc.919

Bronner, I.F., Quail, M.A., Turner, D.J. and Swerdlow, H., 2013. Improved protocols for illumina sequencing. Current protocols in human genetics, 79(1), pp.18-2. https://doi.org/10.1002/0471142905.hg1802s79

Bustin, S., and Huggett, J., 2017. qPCR primer design revisited. Biomolecular detection and quantification, 14, pp.19-28. https://doi.org/10.1016/j.bdq.2017.11.001

Buxton, A.S., Groombridge, J.J. and Griffiths, R.A., 2018. Comparison of two citizen scientist methods for collecting pond water samples for environmental DNA studies. Ctizen Science: Theory and Practice, 3(2), p.2. http://doi.org/10.5334/cstp.151

Bylemans, J., Furlan, E.M., Hardy, C.M., McGuffie, P., Lintermans, M. and Gleeson, D.M., 2017. An environmental DNA-based method for monitoring spawning activity: A case study, using the endangered Macquarie perch (*Macquaria australasica*). Methods in Ecology and Evolution, 8(5), pp.646-655. https://doi.org/10.1111/2041-210X.12709

Bylemans, J., Gleeson, D.M., Hardy, C.M. and Furlan, E., 2018. Toward an ecoregion scale evaluation of eDNA metabarcoding primers: A case study for the freshwater fish biodiversity of the Murray–Darling Basin (Australia). Ecology and Evolution, 8(17), pp.8697-8712. https://doi.org/10.1002/ece3.4387

Bylemans, J., Gleeson, D.M., Duncan, R.P., Hardy, C.M. and Furlan, E.M., 2019. A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. Environmental DNA, 1(4), pp.402-414. https://doi.org/10.1002/edn3.41

Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., Taberlet, P., Vigouroux, R. and Brosse, S., 2019. Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. Scientific reports, 9(1), pp.1-11. DOI: 10.1038/s41598-019-39399-5

Capo, E., Spong, G., Norman, S., Königsson, H., Bartels, P. and Byström, P., 2019. Droplet digital PCR assays for the quantification of brown trout (*Salmo trutta*) and Arctic char (*Salvelinus alpinus*) from environmental DNA collected in the water of mountain lakes. PloS one, 14(12), p.e0226638. https://doi.org/10.1371/journal.pone.0226638

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., Iribar, A., Taberlet, P., Vigouroux, R., Brosse, S., 2019. Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. Molecular Ecology Resources, 19(1), pp.27-46. https://doi.org/10.1111/1755-0998.12900

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J.C., Bonin, A., Taberlet, P. and Pont, D., 2016. Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. PloS one, 11(6), p.e0157366. doi: 10.1371/journal.pone.0157366

Clarke, R.T. and Murphy, J.F., 2006. Effects of locally rare taxa on the precision and sensitivity of RIVPACS bioassessment of freshwaters. Freshwater Biology, 51(10), pp.1924-1940. https://doi.org/10.1111/j.1365-2427.2006.01611.x

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

Collen, B.E.N., Loh, J., Whitmee, S., McRae, L., Amin, R. and Baillie, J.E., 2009. Monitoring change in vertebrate abundance: the Living Planet Index. Conservation Biology, 23(2), pp.317-327. https://doi.org/10.1111/j.1523-1739.2008.01117.xRees

Collen, B., Pettorelli, N., Baillie, J.E. and Durant, S.M., 2013. Biodiversity monitoring and conservation: bridging the gap between global commitment and local action. Biodiversity monitoring and conservation: Bridging the gaps between global commitment and local action, pp.1-16.

Collins, R. A., Wangensteen, O. S., O'Gorman, E. J., Mariani, S., Sims, D. W., Genner, M. J., 2018. Persistence of environmental DNA in marine systems. Communications Biology, 1(1), pp.1-11. https://doi.org/10.1038/s42003-018-0192-6

Collins, R.A., Bakker, J., Wangensteen, O.S., Soto, A.Z., Corrigan, L., Sims, D.W., Genner, M.J. and Mariani, S., 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution, 10(11), pp.1985-2001. https://doi.org/10.1111/2041-210X.13276

Cooke, S.J., Paukert, C. and Hogan, Z., 2012. Endangered river fish: factors hindering conservation and restoration. Endangered species research, 17(2), pp.179-191. https://doi.org/10.3354/esr00426

Cooke, S.J., Hogan, Z.S., Butcher, P.A., Stokesbury, M.J., Raghavan, R., Gallagher, A.J., Hammerschlag, N. and Danylchuk, A.J., 2016. Angling for endangered fish: conservation problem or conservation action?. Fish and Fisheries, 17(1), pp.249-265. https://doi.org/10.1111/faf.12076

Copp, G.H., Wesley, K.J., Verreycken, H. and Russell, I.C., 2007. When an 'invasive' fish species fails to invade! Example of the topmouth gudgeon *Pseudorasbora parva*. Aquatic Invasions, 2(2), pp.107-112. doi: 10.3391/ai.2007.2.2.4

Corrigan, L.J., Lucas, M.C., Winfield, I.J. and Hoelzel, A.R., 2011. Environmental factors associated with genetic and phenotypic divergence among sympatric populations of Arctic charr (*Salvelinus alpinus*). Journal of Evolutionary Biology, 24(9), pp.1906-1917. https://doi.org/10.1111/j.1420-9101.2011.02327.x

Cowart, D.A., Breedveld, K.G., Ellis, M.J., Hull, J.M. and Larson, E.R., 2018. Environmental DNA (eDNA) applications for the conservation of imperiled crayfish (Decapoda: Astacidea) through monitoring of invasive species barriers and relocated populations. Journal of Crustacean Biology, 38(3), pp.257-266. https://doi.org/10.1093/jcbiol/ruy007

Cristescu, M. E. and Hebert, P. D., 2018. Uses and misuses of environmental DNA in biodiversity science and conservation. Annual Review of Ecology, Evolution, and Systematics, 49, 209-230. https://doi.org/10.1146/annurev-ecolsys-110617- 062306

Davison, P.I., Copp, G.H., Créach, V., Vilizzi, L. and Britton, J.R., 2017. Application of environmental DNA analysis to inform invasive fish eradication operations. The Science of Nature, 104(3-4), p.35. DOI: 10.1007/s00114-017-1453-9.

Davison, P.I., Falcou-Préfol, M., Copp, G.H., Davies, G.D., Vilizzi, L. and Créach, V., 2019. Is it absent or is it present? Detection of a non-native fish to inform management decisions using a new highly-sensitive eDNA protocol. Biological Invasions, 21(8), pp.2549-2560. DOI: 10.1007/s10530-019-01993-z

Deagle, B.E., Thomas, A.C., Shaffer, A.K., Trites, A.W. and Jarman, S.N., 2013. Quantifying sequence proportions in a DNA-based diet study using Ion Torrent

amplicon sequencing: which counts count?. Molecular Ecology Resources, 13(4), pp.620-633. https://doi.org/10.1111/1755-0998.12103

Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R. and Eveson, J.P., 2019. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data?. Molecular ecology, 28(2), pp.391-406. https://doi.org/10.1111/mec.14734

Deakin, C.T., Deakin, J.J., Ginn, S.L., Young, P., Humphreys, D., Suter, C.M., Alexander, I.E. and Hallwirth, C.V., 2014. Impact of next-generation sequencing error on analysis of barcoded plasmid libraries of known complexity and sequence. Nucleic acids research, 42(16), pp.e129-e129. https://doi.org/10.1093/nar/gku607

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E. and Taberlet, P., 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. Molecular ecology resources, 14(2), pp.306-323. DOI: 10.1111/1755-0998.12188

Deiner, K., and Altermatt, F., 2014. Transport distance of invertebrate environmental DNA in a natural river. PloS one, 9(2), e88786. https://doi.org/10.1371/journal.pone.0088786

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

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., De Vere, N. and Pfrender, M.E., 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular ecology, 26(21), pp.5872-5895. DOI: 10.1111/mec.14350

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P. and Miaud, C., 2011. Persistence of environmental DNA in freshwater ecosystems. PloS one, 6(8), p.e23398. https://doi.org/10.1371/journal.pone.0023398

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E. and Miaud, C., 2012. Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. Journal of applied ecology, 49(4), pp.953-959. https://doi.org/10.1111/j.1365-2664.2012.02171.x

Delaney, D.G. and Leung, B., 2010. An empirical probability model of detecting species at low densities. Ecological Applications, 20(4), pp.1162-1172. https://doi.org/10.1890/09-0309.1

Dennerline, D.E., Jennings, C.A. and Degan, D.J., 2012. Relationships between hydroacoustic derived density and gill net catch: implications for fish assessments. Fisheries Research, 123, pp.78-89. https://doi.org/10.1016/j.fishres.2011.11.012

De Souza, L.S., Godwin, J.C., Renshaw, M.A. and Larson, E., 2016. Environmental DNA (eDNA) detection probability is influenced by seasonal activity of organisms. PloS one, 11(10), p.e0165273. DOI: 10.1371/journal.pone.0165273

Djurhuus, A., Port, J., Closek, C.J., Yamahara, K.M., Romero-Maraccini, O., Walz, K.R., Goldsmith, D.B., Michisaki, R., Breitbart, M., Boehm, A.B. and Chavez, F.P., 2017. Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. Frontiers in Marine Science, 4, p.314. https://doi.org/10.3389/fmars.2017.00314

Doble, C. J., Hipperson, H., Salzburger, W., Horsburgh, G. J., Mwita, C., Murrell, D. J., Day, J. J., 2020. Testing the performance of environmental DNA metabarcoding for surveying highly diverse tropical fish communities: A case study from Lake Tanganyika. Environmental DNA, 2(1), pp.24-41. https://doi.org/10.1002/edn3.43

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T. and Minamoto, T., 2017. Environmental DNA analysis for estimating the abundance and biomass of stream fish. Freshwater Biology, 62(1), pp.30-39. https://doi.org/10.1111/fwb.12846

Doi, H., Fukaya, K., Oka, S.I., Sato, K., Kondoh, M. and Miya, M., 2019. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental

DNA metabarcoding using a multispecies site occupancy model. Scientific reports, 9(1), pp.1-8. https://doi.org/10.1371/journal.pone.0210357

Dorazio, R.M. and Erickson, R.A., 2018. ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. Molecular ecology resources, 18(2), pp.368-380. https://doi.org/10.1111/1755-0998.12735

Dougherty, M.M., Larson, E.R., Renshaw, M.A., Gantz, C.A., Egan, S.P., Erickson, D.M. and Lodge, D.M., 2016. Environmental DNA (eDNA) detects the invasive rusty crayfish *Orconectes rusticus* at low abundances. Journal of Applied Ecology, 53(3), pp.722-732.

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. and Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews, 81(2), pp.163-182. DOI:10.1017/S1464793105006950

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. and Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics, 27(16), pp.2194-2200. https://doi.org/10.1093/bioinformatics/btr381

Eichmiller, J.J., Miller, L.M. and Sorensen, P.W., 2016. Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. Molecular ecology resources, 16(1), pp.56-68. DOI: 10.1111/1755-0998.12421

Eiler, A., Löfgren, A., Hjerne, O., Nordén, S. and Saetre, P., 2018. Environmental DNA (eDNA) detects the pool frog (*Pelophylax lessonae*) at times when traditional monitoring methods are insensitive. Scientific reports, 8(1), pp.1-9. https://doi.org/10.1038/s41598-018-23740-5

Elbrecht, V. and 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), p.e0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., Vamos, E.E., Steinke, D. and Leese, F., 2018. Estimating intraspecific genetic diversity from community DNA metabarcoding data. PeerJ, 6, p.e4644. https://doi.org/10.7717/peerj.4644.

Elliott, J.A. and Bell, V.A., 2011. Predicting the potential long-term influence of climate change on vendace (*Coregonus albula*) habitat in Bassenthwaite Lake, UK. Freshwater Biology, 56(2), pp.395-405. https://doi.org/10.1111/j.1365-2427.2010.02506.x

Eloranta, A.P., Siwertsson, A., Knudsen, R. and Amundsen, P.A., 2011. Dietary plasticity of Arctic charr (*Salvelinus alpinus*) facilitates coexistence with competitively superior European whitefish (*Coregonus lavaretus*). Ecology of Freshwater Fish, 20(4), pp.558-568. https://doi.org/10.1111/j.1600-0633.2011.00504.x

Erickson, R.A., Rees, C.B., Coulter, A.A., Merkes, C.M., McCalla, S.G., Touzinsky, K.F., Walleser, L., Goforth, R.R. and Amberg, J.J., 2016. Detecting the movement and spawning activity of bigheaded carps with environmental DNA. Molecular Ecology Resources, 16(4), pp.957-965. DOI: 10.1111/1755-0998.12533

Erisman, B., Heyman, W., Kobara, S., Ezer, T., Pittman, S., Aburto-Oropeza, O. and Nemeth, R.S., 2017. Fish spawning aggregations: where well-placed management actions can yield big benefits for fisheries and conservation. Fish and Fisheries, 18(1), pp.128-144. https://doi.org/10.1111/faf.12132

Esteve, M., 2005. Observations of spawning behaviour in Salmoninae: *Salmo*, *Oncorhynchus* and *Salvelinus*. Reviews in Fish Biology and Fisheries, 15(1-2), pp.1-21. https://doi.org/10.1007/s11160-005-7434-7

European Communities, 1992. Directive 92/43/EEC, on the conservation of natural habitats and of wild fauna and flora. Official Journal of the European Communities L 206, 7–50.

European Communities, 2000. Directive 2000/60/EC, Establishing a framework for community action in the field of water policy. Official Journal of the European Communities L 327, 1–71.

Eva, B., Harmony, P., Thomas, G., Francois, G., Alice, V., Claude, M. and Tony, D., 2016. Trails of river monsters: Detecting critically endangered Mekong giant catfish *Pangasianodon gigas* using environmental DNA. Global ecology and conservation, 7, pp.148-156. https://doi.org/10.1016/j.gecco.2016.06.007

Evans, N.T., Olds, B.P., Renshaw, M.A., Turner, C.R., Li, Y., Jerde, C.L., Mahon, A.R., Pfrender, M.E., Lamberti, G.A. and Lodge, D.M., 2016. Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Molecular ecology resources, 16(1), pp.29-41. DOI: 10.1111/1755-0998.12433

Evans, N.T., Shirey, P.D., Wieringa, J.G., Mahon, A.R. and Lamberti, G.A., 2017. Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. Fisheries, 42(2), pp.90-99. https://doi.org/10.1080/03632415.2017.1276329

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

Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P. and Pompanon, F., 2010. An *in silico* approach for the evaluation of DNA barcodes. BMC genomics, 11(1), p.434. https://doi.org/10.1186/1471-2164-11-434

Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L., Lopes, C.M., Boyer, F., Pompanon, F. and Rayé, G., 2015. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular ecology resources, 15(3), pp.543-556. https://doi.org/10.1111/1755-0998.12338

Ficetola, G.F., Romano, A., Salvidio, S. and Sindaco, R., 2018. Optimizing monitoring schemes to detect trends in abundance over broad scales. Animal Conservation, 21(3), pp.221-231. https://doi.org/10.1111/acv.12356.

Fonseca, V.G., 2018. Pitfalls in relative abundance estimation using eDNA metabarcoding. Molecular Ecology Resources, 18(5), pp.923-926. https://doi.org/10.1111/1755-0998.12902

Franssen, J., Blais, C., Lapointe, M., Bérubé, F., Bergeron, N., Magnan, P., 2012. Asphyxiation and entombment mechanisms in fines rich spawning substrates: experimental evidence with brook trout (*Salvelinus fontinalis*) embryos. Canadian journal of fisheries and aquatic sciences, 69(3), pp.587-599. https://doi.org/10.1139/f2011-168

Freyhof, J., Kottelat, M., 2005. *Salvelinus evasus* sp. n., a charr from deep waters of Lake Ammersee, southern Germany (Teleostei: Salmonidae), with comments on two extinct species. Revue suisse de Zoologie, 112(1), pp.253-269.

Frost, W.E., 1952. Predators on the eggs of char in Windermere. John Sherratt and Son.

Frost, W.E., 1965. Breeding habits of Windermere charr, *Salvelinus willughbii* (Günther), and their bearing on speciation of these fish. Proceedings of the Royal Society of London. Series B. Biological Sciences, 163(991), pp.232-284. https://doi.org/10.1098/rspb.1965.0070

Frost, W.E., 1977. The food of Charr,* *Salvelinus willughbii* (Günther), in Windermere. Journal of Fish Biology, 11(6), pp.531-547. https://doi.org/10.1111/j.1095-8649.1977.tb05710.x

Furlan, E.M., Gleeson, D., Hardy, C.M. and Duncan, R.P., 2016. A framework for estimating the sensitivity of eDNA surveys. Molecular ecology resources, 16(3), pp.641-654. https://doi.org/10.1111/1755-0998.12483

Garduño-Paz, M.V., Demetriou, M. and Adams, C.E., 2010. Variation in scale shape among alternative sympatric phenotypes of Arctic charr *Salvelinus alpinus* from two lakes in Scotland. Journal of fish biology, 76(6), pp.1491-1497. https://doi.org/10.1111/j.1095-8649.2010.02584.x

Gaudin, P. and Caillere, L., 2000. Experimental study of the influence of presence and predation by sculpin, *Cottus gobio* L., on the drift of emergent brown trout, *Salmo trutta* L. Arch. Hydrobiol, 147(3), pp.257-271.Hydrobiol, 147(3), 257-271.

GB Non-Native Species Secretariat, 2020. www.nonnativespecies.org [Accessed May 2020].

Genovesi, P., and Carnevali, L., 2011. Island invasives: Eradication and man-agement. Gland, Switzerland: IUCN.

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), pp.1333-1345. https://doi.org/10.1111/1755-0998.13201

Girard, M., Goulon, C., Tessier, A., Vonlanthen, P., Guillard, J., 2020. Comparisons of day-time and night-time hydroacoustic surveys in temperate lakes. Aquatic Living Resources, 33, 9. https://doi.org/10.1051/alr/2020011

Gomez-Uchida, D., Dunphy, K. P., O'connell, M. F., Ruzzante, D. E., 2008. Genetic divergence between sympatric Arctic charr *Salvelinus alpinus* morphs in Gander Lake, Newfoundland: roles of migration, mutation and unequal effective population sizes. Journal of Fish Biology, 73(8), pp.2040-2057. https://doi.org/10.1111/j.1095-8649.2008.02048.x

Goldberg, C.S., Sepulveda, A., Ray, A., Baumgardt, J. and Waits, L.P., 2013. Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). Freshwater Science, 32(3), pp.792-800. https://doi.org/10.1899/13-046.1

Goldberg, C.S., Turner, C.R., Deiner, K., Klymus, K.E., Thomsen, P.F., Murphy, M.A., Spear, S.F., McKee, A., Oyler-McCance, S.J., Cornman, R.S. and Laramie, M.B., 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in ecology and evolution, 7(11), pp.1299-1307. https://doi.org/10.1111/2041-210X.12595

Goodwin, S., McPherson, J. D., McCombie, W. R., 2016. Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17(6), 333. https://doi.org/10.1038/nrg.2016.49

Graham, C. T., and Harrod, C., 2009. Implications of climate change for the fishes of the British Isles. Journal of Fish Biology, 74(6), pp.1143-1205. ttps://doi.org/10.1111/j.1095-8649.2009.02180.x

Grey, E.K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K.L., Lacoursière-Roussel, A., Leong, S.C.Y., Li, Y., Olds, B. and Pfrender, M.E., 2018. Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports, 8(1), pp.1-10. https://doi.org/10.1038/s41598-018-27048-2

Griffith, D.M., Veech, J.A. and Marsh, C.J., 2016. cooccur: probabilistic species co-occurrence analysis in R. Journal of Statistical Software, 69(2), pp.1-17. DOI: 10.18637/jss.v069.c02

Griffiths, N.P., Bolland, J.D., Wright, R.M., Murphy, L.A., Donnelly, R.K., Watson, H.V. and Hänfling, B., 2020. Environmental DNA metabarcoding provides enhanced detection of the European eel *Anguilla anguilla* and fish community structure in pumped river catchments. Journal of Fish Biology. https://doi.org/10.1111/jfb.14497

Groot, C., 1996. Salmonid life histories. Dev Aquacult Fish Sci. 29:97–230.

Hänfling, B., Lawson Handley, L., Read, D.S., Hahn, C., Li, J., Nichols, P., Blackman, R.C., Oliver, A. and Winfield, I.J., 2016. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular ecology, 25(13), pp.3101-3119. https://doi.org/10.1111/mec.13660

Hansen, M.J., Krueger, C.C., Muir, A.M., Klemetsen, A. and Power, M., 2019. Assessing the impact of charr research past, present, and future. Hydrobiologia, 840(1), pp.1-10. https://doi.org/10.1007/s10750-019-04012-3

Hansen, B. K., Jacobsen, M. W., Middelboe, A. L., Preston, C. M., Marin, R., Bekkevold, D., Knudsen, S.W., Møller, P.R., Nielsen, E. E., 2020. Remote, autonomous real-time

monitoring of environmental DNA from commercial fish. Scientific reports, 10(1), pp.1-8. https://doi.org/10.1038/s41598-020-70206-8

Hardie, S.A., Barmuta, L.A. and White, R.W., 2006. Comparison of day and night fyke netting, electrofishing and snorkelling for monitoring a population of the threatened golden galaxias (*Galaxias auratus*). Hydrobiologia, 560(1), pp.145-158. https://doi.org/10.1007/s10750-005-9509-9

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. and 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), pp.6330-6341. https://doi.org/10.1002/ece3.4013

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. and Priestley, V., 2019. Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. Hydrobiologia, 826(1), pp.25-41. https://doi.org/10.1007/s10750-018-3750-5

Harper, L.R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H.C., Lewis, E., Adams, I.P., Brotherton, P., Phillips, S. and Hänfling, B., 2020. Generating and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: a case study on a threatened amphibian. Environmental DNA, 2(2), pp.184-199. https://doi.org/10.1002/edn3.57

Hempel, C.A., Peinert, B., Beermann, A.J., Elbrecht, V., Macher, J.N., Macher, T.H., Jacobs, G. and Leese, F., 2020. Using environmental DNA to monitor the reintroduction success of the Rhine sculpin (*Cottus rhenanus*) in a restored stream. Frontiers in Ecology and Evolution, 8, p.81. https://doi.org/10.3389/fevo.2020.00081

Hering, D., Borja, A., Jones, J.I., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S., Hänfling, B., Kahlert, M. and Leese, F., 2018. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water Research, 138, pp.192-205. https://doi.org/10.1016/j.watres.2018.03.003

Hinlo, R., Furlan, E., Suitor, L. and Gleeson, D., 2017. Environmental DNA monitoring and management of invasive fish: comparison of eDNA and fyke netting. Management of Biological Invasions, 8(1), p.89. https://doi.org/10.3391/mbi.2017.8.1.09

Holman, L.E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J. and Rius, M., 2019. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific reports, 9(1), pp.1-10. https://doi.org/10.1038/s41598-019-47899-7

Hunter, M.E., Meigs-Friend, G., Ferrante, J.A., Kamla, A.T., Dorazio, R.M., Diagne, L.K., Luna, F., Lanyon, J.M. and Reid, J.P., 2018. Surveys of environmental DNA (eDNA): a new approach to estimate occurrence in Vulnerable manatee populations. Endangered Species Research, 35, pp.101-111. https://doi.org/10.3354/esr00880

Igoe, F., O'Grady, M.F., Tierney, D. and Fitzmaurice, P., 2003, July. Arctic char *Salvelinus alpinus* (L.) in Ireland: A millennium review of its distribution and status with conservation recommendations. In Biology and Environment: Proceedings of the Royal Irish Academy (pp. 9-22). Royal Irish Academy. https://www.jstor.org/stable/20500174

Itakura, H., Wakiya, R., Yamamoto, S., Kaifu, K., Sato, T. and Minamoto, T., 2019. Environmental DNA analysis reveals the spatial distribution, abundance, and biomass of Japanese eels at the river-basin scale. Aquatic Conservation: Marine and Freshwater Ecosystems, 29(3), pp.361-373. https://doi.org/10.1002/aqc.3058

IUCN FFSG, 2015. Importance of freshwater fishes. Available at: https://www.iucnffsg.org/freshwater-fishes/importance-of-freshwater-fishes/

IUCN's response to the Post-2020 Global Biodiversity Framework discussion paper: Part 2- Target formulations and topics. https://www.iucn.org/sites/dev/files/iucn_response_cbd_post_2020_part_2_target_formulations_and_topics_12_april_2019_final.pdf

Jane, S.F., Wilcox, T.M., McKelvey, K.S., Young, M.K., Schwartz, M.K., Lowe, W.H., Letcher, B.H. and Whiteley, A.R., 2015. Distance, flow and PCR inhibition: e DNA

dynamics in two headwater streams. Molecular ecology resources, 15(1), pp.216-227. https://doi.org/10.1111/1755-0998.12285

Janosik, A.M. and Johnston, C.E., 2015. Environmental DNA as an effective tool for detection of imperiled fishes. Environmental biology of fishes, 98(8), pp.1889-1893. https://doi.org/10.1007/s10641-015-0405-5

Jerde, C.L., Mahon, A.R., Chadderton, W.L. and Lodge, D.M., 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters, 4(2), pp.150-157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Jerde, C.L., Chadderton, W.L., Mahon, A.R., Renshaw, M.A., Corush, J., Budny, M.L., Mysorekar, S. and Lodge, D.M., 2013. Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences, 70(4), pp.522-526. https://doi.org/10.1139/cjfas-2012-0478

Jerde, C.L., 2019. Can we manage fisheries with the inherent uncertainty from eDNA?. Journal of fish biology. https://doi.org/10.1111/jfb.14218

Jo, T., Murakami, H., Yamamoto, S., Masuda, R. and Minamoto, T., 2019. Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. Ecology and evolution, 9(3), pp.1135-1146. https://doi.org/10.1002/ece3.4802

Johnson, L., 1980. The arctic charr, *Salvelinus alpinus*. Charrs, salmonid fishes of the genus *Salvelinus*.

Jonsson, B., and Jonsson, N., 2001. Polymorphism and speciation in Arctic charr. Journal of Fish Biology, 58(3), pp.605-638. https://doi.org/10.1111/j.1095-8649.2001.tb00518.x

Joseph, L.N., Field, S.A., Wilcox, C. and Possingham, H.P., 2006. Presence–absence versus abundance data for monitoring threatened species. Conservation biology, 20(6), pp.1679-1687. https://doi.org/10.1111/j.1523-1739.2006.00529.x

Kamoroff, C. and Goldberg, C.S., 2018. An issue of life or death: using eDNA to detect viable individuals in wilderness restoration. Freshwater Science, 37(3), pp.685-696. https://doi.org/10.1086/699203

Kassambara, A., 2018. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2. https://CRAN.R-project.org/package=ggpubr

Kebschull, J.M. and Zador, A.M., 2015. Sources of PCR-induced distortions in high-throughput sequencing data sets. Nucleic acids research, 43(21), pp.e143-e143. https://doi.org/10.1093/nar/gkv717

Kelly, R.P., Port, J.A., Yamahara, K.M. and Crowder, L.B., 2014. Using environmental DNA to census marine fishes in a large mesocosm. PloS one, 9(1), p.e86175. https://doi.org/10.1371/journal.pone.0086175

Kelly, R.P., Shelton, A.O. and Gallego, R., 2019. Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. Scientific reports, 9(1), pp.1-14. https://doi.org/10.1038/s41598-019-48546-x

Kelly, S., Moore, T. N., de Eyto, E., Dillane, M., Goulon, C., Guillard, J., Lasne, E., McGinnity, P., Poole, R., Winfield, I.J., Woolway, R.I., Jennings, E. (2020). Warming winters threaten peripheral Arctic charr populations of Europe. Climatic Change, 163(1), pp.599-618. https://doi.org/10.1007/s10584-020-02887-z

King, A.J., Ward, K.A., O'connor, P., Green, D., Tonkin, Z. and Mahoney, J., 2010. Adaptive management of an environmental watering event to enhance native fish spawning and recruitment. Freshwater biology, 55(1), pp.17-31. https://doi.org/10.1111/j.1365-2427.2009.02178.x

Klemetsen, A., Amundsen, P.A., Dempson, J.B., Jonsson, B., Jonsson, N., O'connell, M.F. and Mortensen, E., 2003. Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): a review of aspects of their life histories. Ecology of freshwater fish, 12(1), pp.1-59. https://doi.org/10.1034/j.1600-0633.2003.00010.x

Klemetsen, A., Knudsen, R., Primicerio, R., & Amundsen, P. A., 2006. Divergent, genetically based feeding behaviour of two sympatric Arctic charr, *Salvelinus alpinus* (L.), morphs. Ecology of Freshwater Fish, 15(3), pp.350-355. https://doi.org/10.1111/j.1600-0633.2006.00128.x

Klemetsen, A., 2010. The charr problem revisited: exceptional phenotypic plasticity promotes ecological speciation in postglacial lakes. Freshwater Reviews, 3(1), pp.49-74. https://doi.org/10.1608/FRJ-3.1.3

Klobucar, S.L., Rodgers, T.W. and Budy, P., 2017. At the forefront: evidence of the applicability of using environmental DNA to quantify the abundance of fish populations in natural lentic waters with additional sampling considerations. Canadian Journal of Fisheries and Aquatic Sciences, 74(12), pp.2030-2034. https://doi.org/10.1139/cjfas-2017-0114

Klymus, K.E., Richter, C.A., Chapman, D.C. and Paukert, C., 2015. Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. Biological Conservation, 183, pp.77-84.http://dx.doi.org/10.1016/j.biocon.2014.11.020

Klymus, K.E., Marshall, N.T. and Stepien, C.A., 2017. Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. PLoS One, 12(5), p.e0177643. https://doi.org/10.1371/journal.pone.0177643

Koenig, C.C., Coleman, F.C., Grimes, C.B., Fitzhugh, G.R., Scanlon, K.M., Gledhill, C.T. and Grace, M., 2000. Protection of fish spawning habitat for the conservation of warm-temperate reef-fish fisheries of shelf-edge reefs of Florida. Bulletin of Marine Science, 66(3), pp.593-616.

Kull, T., Sammul, M., Kull, K., Lanno, K., Tali, K., Gruber, B., Schmeller, D. and Henle, K., 2008. Necessity and reality of monitoring threatened European vascular plants. Biodiversity and Conservation, 17(14), p.3383. DOI: 10.1007/s10531-008-9432-2

Lacoursière-Roussel, A., Côté, G., Leclerc, V. and Bernatchez, L., 2016a. Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. Journal of Applied Ecology, 53(4), pp.1148-1157. https://doi.org/10.1111/1365-2664.12598

Lacoursière-Roussel, A., Rosabal, M. and Bernatchez, L., 2016b. Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. Molecular ecology resources, 16(6), pp.1401-1414. DOI: 10.1111/1755-0998.12522

Lacoursière-Roussel, A., Dubois, Y., Normandeau, E. and Bernatchez, L., 2016c. Improving herpetological surveys in eastern North America using the environmental DNA method. Genome, 59(11), pp.991-1007. https://doi.org/10.1139/gen-2015-0218

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

Laramie, M.B., Pilliod, D.S. and Goldberg, C.S., 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biological Conservation, 183, pp.29-37. https://doi.org/10.1016/j.biocon.2014.11.025

Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics, 30(22), pp.3276-3278. https://doi.org/10.1093/bioinformatics/btu531

Larson, E.R., Graham, B.M., Achury, R., Coon, J.J., Daniels, M.K., Gambrell, D.K., Jonasen, K.L., King, G.D., LaRacuente, N., Perrin-Stowe, T.I. and Reed, E.M., 2020. From eDNA to citizen science: emerging tools for the early detection of invasive species. Frontiers in Ecology and the Environment, 18(4), pp.194-202. https://doi.org/10.1002/fee.2162

Lawson Handley, L., 2015. How will the 'molecular revolution' contribute to biological recording?. Biological Journal of the Linnean Society, 115(3), pp.750-766. https://doi.org/10.1111/bij.12516

Lawson Handley, L., Read, D.S., Winfield, I.J., Kimbell, H., Johnson, H., Li, J., Hahn, C., Blackman, R., Wilcox, R., Donnelly, R. and Szitenberg, A., 2019. Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. Environmental DNA, 1(1), pp.26-39. https://doi.org/10.1002/edn3.5

Lecaudey, L. A., Schletterer, M., Kuzovlev, V. V., Hahn, C., Weiss, S. J., 2019. Fish diversity assessment in the headwaters of the Volga River using environmental DNA metabarcoding. Aquatic Conservation: Marine and Freshwater Ecosystems, 29(10), pp.1785-1800. https://doi.org/10.1002/aqc.3163

Levasseur, M., Bergeron, N. E., Lapointe, M. F., Bérubé, F., 2006. Effects of silt and very fine sand dynamics in Atlantic salmon (*Salmo salar*) redds on embryo hatching success. Canadian Journal of Fisheries and Aquatic Sciences, 63(7), pp.1450-1459. https://doi.org/10.1139/f06-050

Levêque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J. and Tedesco, P.A., 2007. Global diversity of fish (Pisces) in freshwater. In Freshwater animal diversity assessment (pp. 545-567). Springer, Dordrecht. https://doi.org/10.1007/s10750-007 -903

Li, J., Lawson Handley, L.J., Read, D.S. and Hänfling, B., 2018. The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. Molecular Ecology Resources, 18(5), pp.1102-1114. https://doi.org/10.1111/1755-0998.12899

Li, J., Hatton-Ellis, T.W., Lawson Handley, L.J., Kimbell, H.S., Benucci, M., Peirson, G. and Hänfling, B., 2019a. Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. Journal of Applied Ecology, 56(5), pp.1232-1244. https://doi.org/10.1111/1365-2664.13352

Li, J., Lawson Handley, L.J., Harper, L.R., Brys, R., Watson, H.V., Di Muri, C., Zhang, X. and Hänfling, B., 2019b. Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. Environmental DNA, 1(3), pp.238-250.https://doi.org/10.1002/edn3.24

Lintermans, M., 2016. Finding the needle in the haystack: comparing sampling methods for detecting an endangered freshwater fish. Marine and Freshwater Research, 67(11), pp.1740-1749. https://doi.org/10.1071/MF14346

Lodge, D.M., Turner, C.R., Jerde, C.L., Barnes, M.A., Chadderton, L., Egan, S.P., Feder, J.L., Mahon, A.R. and Pfrender, M.E., 2012. Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. Molecular ecology, 21(11), pp.2555-2558. https://doi.org/10.1111/j.1365-294X.2012.05600.x

Low, J.J., Igoe, F., Davenport, J. and Harrison, S.S., 2011. Littoral spawning habitats of three southern Arctic charr (*Salvelinus alpinus* L.) populations. Ecology of Freshwater Fish, 20(4), pp.537-547. https://doi.org/10.1111/j.1600-0633.2011.00502.x

Lyle, A., Maitland, P.S., Winfield, I.J., 2006. Translocation of vendace from Derwentwater to safe refuge locations 2005/2066. Final Report.

Lyle, A.A., and Adams, C.E., 2017. Assessing the presence of vendace in Loch Earn, 2016. Scottish Natural Heritage Commissioned Report No. 976.

Lyle, A.A., Bean, C.W., Honkanen, H.M., Koene, J.P. and Adams, C.E., 2019. The establishment of Scotland's rarest freshwater fish, the vendace (*Coregonus albula*), in conservation refuge sites. Glasgow Naturalist (online 2019), 27(part 1).

Lynch, A.J., Cooke, S.J., Deines, A.M., Bower, S.D., Bunnell, D.B., Cowx, I.G., Nguyen, V.M., Nohner, J., Phouthavong, K., Riley, B. and Rogers, M.W., 2016. The social, economic, and environmental importance of inland fish and fisheries. Environmental Reviews, 24(2), pp.115-121. https://doi.org/10.1139/er-2015-0064

MacKenzie, B.R., Tomkiewicz, J., Köster, F. and Nissling, A., 1998. Quantifying and disaggregating the spawner effect: Incorporating stock structure, spatial distribution and female influences into estimates of annual population egg production.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J. and Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology, 83(8), pp.2248-2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

MacKenzie, D.I. and Nichols, J.D., 2004. Occupancy as a surrogate for abundance estimation. Animal biodiversity and conservation, 27(1), pp.461-467. https://www.raco.cat/index.php/ABC/article/view/57325

Magoč, T. and Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics, 27(21), pp.2957-2963.

Mahon, A.R., Jerde, C.L., Galaska, M., Bergner, J.L., Chadderton, W.L., Lodge, D.M., Hunter, M.E. and Nico, L.G., 2013. Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. PloS one, 8(3), p.e58316. https://doi.org/10.1371/journal.pone.0058316

Maitland, P.S., Winfield, I.J., McCarthy, I.D. and Igoe, F., 2007. The status of Arctic charr *Salvelinus alpinus* in Britain and Ireland. Ecology of Freshwater Fish, 16(1), pp.6-19. https://doi.org/10.1111/j.1600-0633.2006.00167.x

Maitland, P.S. and Lyle, A.A., 2013. E x situ and in situ approaches, including assisted reproduction, for the conservation of native species of charr (Salmonidae) and whitefish (Coregonidae) in Scotland. International Zoo Yearbook, 47(1), pp.129-139. https://doi.org/10.1111/j.1748-1090.2012.00192.x

Majaneva, M., Diserud, O.H., Eagle, S.H., Boström, E., Hajibabaei, M. and Ekrem, T., 2018. Environmental DNA filtration techniques affect recovered biodiversity. Scientific reports, 8(1), pp.1-11. DOI: 10.1038/s41598-018-23052-8

Marques, V., Guérin, P.É., Rocle, M., Valentini, A., Manel, S., Mouillot, D. and Dejean, T., 2020. Blind assessment of vertebrate taxonomic diversity across spatial scales by clustering environmental DNA metabarcoding sequences. Ecography. https://doi.org/10.1111/ecog.05049

Marshall, C.T., Kjesbu, O.S., Yaragina, N.A., Solemdal, P. and Ulltang, Ø., 1998. Is spawner biomass a sensitive measure of the reproductive and recruitment potential of Northeast Arctic cod?. Canadian Journal of Fisheries and Aquatic Sciences, 55(7), pp.1766-1783. https://doi.org/10.1139/f98-062

Marshall, N.T. and 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), pp.3515-3538. https://doi.org/10.1002/ece3.4985

Martellini, A., Payment, P. and Villemur, R., 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water research, 39(4), pp.541-548. https://doi.org/10.1016/j.watres.2004.11.012

Maxwell, D. and Jennings, S., 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. Journal of Applied Ecology, 42(1), pp.25-37. https://doi.org/10.1111/j.1365-2664.2005.01000.x

McCallum, M.L., 2015. Vertebrate biodiversity losses point to a sixth mass extinction. Biodiversity and Conservation, 24(10), pp.2497-2519. https://doi.org/10.1007/s10531-015-0940-6

McClenaghan, B., Compson, Z.G. and Hajibabaei, M., 2020. Validating metabarcoding-based biodiversity assessments with multi-species occupancy models: A case study using coastal marine eDNA. PloS one, 15(3), p.e0224119. https://doi.org/10.1371/journal.pone.0224119

McKee, A.M., Spear, S.F. and Pierson, T.W., 2015. The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. Biological Conservation, 183, pp.70-76. https://doi.org/10.1016/j.biocon.2014.11.031.

McQuillan, J. S., and Robidart, J. C., 2017. Molecular-biological sensing in aquatic environments: recent developments and emerging capabilities. Current opinion in biotechnology, 45, pp.43-50. https://doi.org/10.1016/j.copbio.2016.11.022

Miller, H., Winfield, I.J., Fletcher, J.M., Ben James, J., Van Rijn, J., Bull, J.M. and Cotterill, C.J., 2015. Distribution, characteristics and condition of Arctic charr

(*Salvelinus alpinus*) spawning grounds in a differentially eutrophicated twin-basin lake. Ecology of Freshwater Fish, 24(1), pp.32-43. https://doi.org/10.1111/eff.12122

Mills, C.A., Heaney, S.I., Butterwick, C., Corry, J.E. and Elliott, J.M., 1990. Lake enrichment and the status of Windermere charr, *Salvelinus alpinus* (L.). Journal of Fish Biology, 37, pp.167-174. https://doi.org/10.1111/j.1095-8649.1990.tb05032.x

Minamoto, T., Naka, T., Moji, K. and Maruyama, A., 2016. Techniques for the practical collection of environmental DNA: filter selection, preservation, and extraction. Limnology, 17(1), pp.23-32. https://doi.org/10.1007/s10201-015-0457-4

Mize, E.L., Erickson, R.A., Merkes, C.M., Berndt, N., Bockrath, K., Credico, J., Grueneis, N., Merry, J., Mosel, K., Tuttle-Lau, M. and Von Ruden, K., 2019. Refinement of eDNA as an early monitoring tool at the landscape-level: study design considerations. Ecological Applications, 29(6), p.e01951. https://doi.org/10.1002/eap.1951

Mizumoto, H., Urabe, H., Kanbe, T., Fukushima, M. and Araki, H., 2018. Establishing an environmental DNA method to detect and estimate the biomass of Sakhalin taimen, a critically endangered Asian salmonid. Limnology, 19(2), pp.219-227. https://doi.org/10.1007/s10201-017-0535-x

Morey, K. C., Bartley, T. J., Hanner, R. H., 2020. Validating environmental DNA metabarcoding for marine fishes in diverse ecosystems using a public aquarium. Environmental DNA, 2(3), pp.330-342. https://doi.org/10.1002/edn3.76

Muhlfeld, C. C., Kovach, R. P., Jones, L. A., Al-Chokhachy, R., Boyer, M. C., Leary, R. F., Winsor, H.L., Luikart, G., Allendorf, F. W., 2014. Invasive hybridization in a threatened species is accelerated by climate change. Nature Climate Change, 4(7), pp.620-624. https://doi.org/10.1038/nclimate2252

Murphy, B.R. and Willis, D.W. eds., 1996. Fisheries techniques (2nd ed., p. 732). Bethesda, Maryland: American Fisheries Society.

Nester, G. M., De Brauwer, M., Koziol, A., West, K. M., DiBattista, J. D., White, N. E., Power, M., Heydenrych, M.J., Harvey, E., Bunce, M., 2020. Development and

evaluation of fish eDNA metabarcoding assays facilitate the detection of cryptic seahorse taxa (family: Syngnathidae). Environmental DNA, 2(4), pp.614-626.

Nichols, J.D. and Williams, B.K., 2006. Monitoring for conservation. Trends in ecology & evolution, 21(12), pp.668-673. https://doi.org/10.1016/j.tree.2006.08.007

Ogram, A., Sayler, G.S. and Barkay, T., 1987. The extraction and purification of microbial DNA from sediments. Journal of microbiological methods, 7(2-3), pp.57-66. https://doi.org/10.1016/0167-7012(87)90025-X

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'hara, R.B., Simpson, G.L., Wagner, H., 2010. Vegan: community ecology package. R package version 1: 17-4.

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. and Pfrender, M.E., 2016. Estimating species richness using environmental DNA. Ecology and evolution, 6(12), pp.4214-4226. https://doi.org/10.1002/ece3.2186

OS Open Map - Local [SHAPE geospatial data], Scale 1:10000, Tiles: ny,sd, Updated: 11 October 2019, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, <https://digimap.edina.ac.uk>, Downloaded: 2020-02-14 10:00:11.794

Pall, K. and Moser, V., 2009. Austrian Index Macrophytes (AIM-Module 1) for lakes: a Water Framework Directive compliant assessment system for lakes using aquatic macrophytes. Hydrobiologia, 633(1), p.83. https://doi.org/10.1007/s10750-009-9871-0

Palm, D., Lindberg, M., Brännäs, E., Lundqvist, H., Östergren, J. and Carlsson, U., 2009. Influence of European sculpin, *Cottus gobio*, on Atlantic salmon *Salmo salar*, recruitment and the effect of gravel size on egg predation–implications for spawning habitat restoration. Fisheries Management and Ecology, 16(6), pp.501-507. https://doi.org/10.1111/j.1365-2400.2009.00705.x

Parducci, L., Alsos, I. G., Unneberg, P., Pedersen, M. W., Han, L., Lammers, Y., Sakari Salonen, J., Väliranta, M.M., Slotte, T., Wohlfarth, B., 2019. Shotgun environmental DNA, pollen, and macrofossil analysis of lateglacial lake sediments from southern

Sweden. Frontiers in Ecology and Evolution, 7, 189.
https://doi.org/10.3389/fevo.2019.00189

Parsons, K.M., Everett, M., Dahlheim, M. and Park, L., 2018. Water, water everywhere: environmental DNA can unlock population structure in elusive marine species. Royal Society open science, 5(8), p.180537. https://doi.org/10.1098/rsos.180537

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., Potter, B.A., Nielsen, R., Rainville, R.A., Orlando, L., Meltzer, D., Kjær, K.H., Willerslev, E., 2016. Postglacial viability and colonization in North America's ice-free corridor. Nature, 537(7618), pp.45-49. https://doi.org/10.1038/nature19085

Pfleger, M.O., Rider, S.J., Johnston, C.E. and Janosik, A.M., 2016. Saving the doomed: Using eDNA to aid in detection of rare sturgeon for conservation (Acipenseridae). Global ecology and conservation, 8, pp.99-107. https://doi.org/10.1016/j.gecco.2016.08.008

Piaggio, A.J., Engeman, R.M., Hopken, M.W., Humphrey, J.S., Keacher, K.L., Bruce, W.E. and 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), pp.374-380. https://doi.org/10.1111/1755-0998.12180

Piccolo, J. J., 2017. The Land Ethic and conservation of native salmonids. Ecology of Freshwater Fish, 26(1), pp.160-164. DOI: 10.1111/eff.12263

Piggott, M.P., 2016. Evaluating the effects of laboratory protocols on eDNA detection probability for an endangered freshwater fish. Ecology and evolution, 6(9), pp.2739-2750. https://doi.org/10.1002/ece3.2083

Pilliod, D.S., Goldberg, C.S., Arkle, R.S. and Waits, L.P., 2013. Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. Canadian Journal of Fisheries and Aquatic Sciences, 70(8), pp.1123-1130. https://doi.org/10.1139/cjfas-2013-0047

Pilliod, D.S., Goldberg, C.S., Arkle, R.S. and Waits, L.P., 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. Molecular Ecology Resources, 14(1), pp.109-116. https://doi.org/10.1111/1755-0998.12159

Piñol, J., Senar, M.A. and Symondson, W.O., 2019. The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. Molecular ecology, 28(2), pp.407-419. https://doi.org/10.1111/mec.14776

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., Roset, N., Schabuss, M., Zornig, H. and Dejean, T., 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. Scientific reports, 8(1), pp.1-13. https://doi.org/10.1038/s41598-018-28424-8

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

Poos, M.S., Mandrak, N.E. and McLaughlin, R.L., 2007. The effectiveness of two common sampling methods for assessing imperilled freshwater fishes. Journal of Fish Biology, 70(3), pp.691-708. https://doi.org/10.1111/j.1095-8649.2007.01349.x

Port, J.A., O'Donnell, J.L., Romero-Maraccini, O.C., Leary, P.R., Litvin, S.Y., Nickols, K.J., Yamahara, K.M. and Kelly, R.P., 2016. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Molecular ecology, 25(2), pp.527-541. https://doi.org/10.1111/mec.13481

Power, M., Power, G., Reist, J. D., Bajno, R., 2009. Ecological and genetic differentiation among the Arctic charr of Lake Aigneau, Northern Québec. Ecology of Freshwater Fish, 18(3), pp.445-460. https://doi.org/10.1111/j.1600-0633.2009.00362.x

Preston, C. M., Harris, A., Ryan, J. P., Roman, B., Marin III, R., Jensen, S., Everlove, C., Birch, J., Dzenitis, J.M., Pargett, D., Adachi, M., Turk, K., Zehr, J.P., Scholin, C. A., 2011.

Underwater application of quantitative PCR on an ocean mooring. PLoS one, 6(8), e22522. https://doi.org/10.1371/journal.pone.0022522

Qu, C., and Stewart, K. A., 2019. Evaluating monitoring options for conservation: comparing traditional and environmental DNA tools for a critically endangered mammal. The Science of Nature, 106(3), pp.1-9. https://doi.org/10.1007/s00114-019-1605-1

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Radinger, J., Britton, J.R., Carlson, S.M., Magurran, A.E., Alcaraz-Hernández, J.D., Almodóvar, A., Benejam, L., Fernández-Delgado, C., Nicola, G.G., Oliva-Paterna, F.J. and Torralva, M., 2019. Effective monitoring of freshwater fish. Fish and Fisheries, 20(4), pp.729-747. https://doi.org/10.1111/faf.12373

Ramsbottom, A. E. (1976). Depth charts of the Cumbrian Lakes - Freshwater Biological Association Scientific Publication No. 33.

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R. and Gough, K.C., 2014. The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. Journal of Applied Ecology, 51(5), pp.1450-1459.https://doi.org/10.1111/1365-2664.12306

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P. and Coissac, E., 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Research, 39(21), pp.e145-e145. DOI: 10.1093/nar/gkr732

Riaz, M., Kuemmerlen, M., Wittwer, C., Cocchiararo, B., Khaliq, I., Pfenninger, M. and Nowak, C., 2020. Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. Ecological Applications, 30(2), p.e02034. https://doi.org/10.1002/eap.2034

Ricciardi, A. and Rasmussen, J.B., 1999. Extinction rates of North American freshwater fauna. Conservation biology, 13(5), pp.1220-1222. https://doi.org/10.1046/j.1523-1739.1999.98380.x

Robinson, C.V., Garcia de Leaniz, C., Rolla, M. and Consuegra, S., 2019. Monitoring the eradication of the highly invasive topmouth gudgeon (*Pseudorasbora parva*) using a novel eDNA assay. Environmental DNA, 1(1), pp.74-85. https://doi.org/10.1002/edn3.12

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

Ruppert, K.M., Kline, R.J. and 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. Global Ecology and Conservation, 17, p.e00547. DOI: 10.1016/j.gecco.2019.e00547

Russell, J. A., Campos, B., Stone, J., Blosser, E. M., Burkett-Cadena, N., Jacobs, J. L., 2018. Unbiased strain-typing of arbovirus directly from mosquitoes using nanopore sequencing: a field-forward biosurveillance protocol. Scientific reports, 8(1), pp.1-12. https://doi.org/10.1038/s41598-018-23641-7

Sakata, M.K., Maki, N., Sugiyama, H. and Minamoto, T., 2017. Identifying a breeding habitat of a critically endangered fish, *Acheilognathus typus*, in a natural river in Japan. The Science of Nature, 104(11-12), p.100. https://doi.org/10.1007/s00114-017-1521-1

Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A. and Leemans, R., 2000. Global biodiversity scenarios for the year 2100. science, 287(5459), pp.1770-1774. DOI: 10.1126/science.287.5459.1770

Sales, N. G., Wangensteen, O. S., Carvalho, D. C., Mariani, S., 2019. Influence of preservation methods, sample medium and sampling time on eDNA recovery in a neotropical river. Environmental DNA, 1(2). https://doi.org/10.1002/edn3.14

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. and Ochu, E., 2020. Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities

using eDNA from riverine systems. Journal of Applied Ecology, 57(4), pp.707-716. https://doi.org/10.1111/1365-2664.13592

Salter, I., Joensen, M., Kristiansen, R., Steingrund, P. and Vestergaard, P., 2019. Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. Communications biology, 2(1), pp.1-9. https://doi.org/10.1038/s42003-019-0696-8

Sandlund, O.T., Museth, J., Næsje, T.F., Rognerud, S., Saksgård, R., Hesthagen, T. and Borgstrøm, R., 2010. Habitat use and diet of sympatric Arctic charr (*Salvelinus alpinus*) and whitefish (*Coregonus lavaretus*) in five lakes in southern Norway: not only interspecific population dominance?. Hydrobiologia, 650(1), pp.27-41. https://doi.org/10.1007/s10750-009-0075-4

Sard, N.M., Herbst, S.J., Nathan, L., Uhrig, G., Kanefsky, J., Robinson, J.D. and Scribner, K.T., 2019. Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. Environmental DNA, 1(4), pp.368-384. https://doi.org/10.1002/edn3.38

Sassoubre, L.M., Yamahara, K.M., Gardner, L.D., Block, B.A. and Boehm, A.B., 2016. Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental science & technology, 50(19), pp.10456-10464. DOI: 10.1021/acs.est.6b03114

Sato, H., Sogo, Y., Doi, H. and Yamanaka, H., 2017. Usefulness and limitations of sample pooling for environmental DNA metabarcoding of freshwater fish communities. Scientific Reports, 7(1), pp.1-12. https://doi.org/10.1038/s41598-017-14978-6

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. https://doi.org/10.1002/rra.3610

Schmidt, B.R., Kery, M., Ursenbacher, S., Hyman, O.J. and Collins, J.P., 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: a

case study of an emerging amphibian pathogen. Methods in Ecology and Evolution, 4(7), pp.646-653. https://doi.org/10.1111/2041-210X.12052

Schnell, I.B., Bohmann, K. and Gilbert, M.T.P., 2015. Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. Molecular ecology resources, 15(6), pp.1289-1303. https://doi.org/10.1111/1755-0998.12402

Schultz, M.T. and Lance, R.F., 2015. Modeling the sensitivity of field surveys for detection of environmental DNA (eDNA). PloS one, 10(10), p.e0141503. https://doi.org/10.1371/journal.pone.0141503

Schwartz, K.R., Parsons, E.C.M., Rockwood, L. and Wood, T.C., 2017. Integrating *in-situ* and *ex-situ* data management processes for biodiversity conservation. Frontiers in Ecology and Evolution, 5, p.120. https://doi.org/10.3389/fevo.2017.00120

Seddon, P.J., Armstrong, D.P. and Maloney, R.F., 2007. Developing the science of reintroduction biology. Conservation biology, 21(2), pp.303-312. https://doi.org/10.1111/j.1523-1739.2006.00627.x

Seddon, P.J., Griffiths, C.J., Soorae, P.S. and Armstrong, D.P., 2014. Reversing defaunation: restoring species in a changing world. Science, 345(6195), pp.406-412. DOI: 10.1126/science.1251818

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

Seymour, M., Durance, I., Cosby, B.J., Ransom-Jones, E., Deiner, K., Ormerod, S.J., Colbourne, J.K., Wilgar, G., Carvalho, G.R., de Bruyn, M. and Edwards, F., 2018. Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. Communications biology, 1(1), pp.1-8. DOI: 10.1038/s42003-017-0005-3

Shen, Y., Hubert, N., Huang, Y., Wang, X., Gan, X., Peng, Z., He, S., 2019. DNA barcoding the ichthyofauna of the Yangtze River: Insights from the molecular inventory of a

mega-diverse temperate fauna. Molecular ecology resources, 19(5), pp.1278-1291. https://doi.org/10.1111/1755-0998.12961

Sigsgaard, E.E., Carl, H., Møller, P.R. and Thomsen, P.F., 2015. Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. Biological Conservation, 183, pp.46-52. https://doi.org/10.1016/j.biocon.2014.11.023

Sigsgaard, E.E., Nielsen, I.B., Bach, S.S., Lorenzen, E.D., Robinson, D.P., Knudsen, S.W., Pedersen, M.W., Al Jaidah, M., Orlando, L., Willerslev, E. and Møller, P.R., 2016. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. Nature ecology & evolution, 1(1), pp.1-5. https://doi.org/10.1038/s41559-016-0004

Sigsgaard, E.E., Torquato, F., Frøslev, T.G., Moore, A.B., Sørensen, J.M., Range, P., Ben-Hamadou, R., Bach, S.S., Møller, P.R. and Thomsen, P.F., 2020. Using vertebrate environmental DNA from seawater in biomonitoring of marine habitats. Conservation Biology, 34(3), pp.697-710. DOI: 10.1111/cobi.13437

Smalås, A., Amundsen, P. A., and Knudsen, R., 2013. Contrasting life history strategies of sympatric Arctic charr morphs, *Salvelinus alpinus*. Journal of ichthyology, 53(10), pp.856-866. https://doi.org/10.1134/S0032945213100111

Smith, D.P. and Peay, K.G., 2014. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. PloS one, 9(2), p.e90234. https://doi.org/10.1371/journal.pone.0090234

Smyly, W.J.P., 1957, June. The life-history of the bullhead or miller's thumb (*Cottus gobio* L.). In Proceedings of the Zoological Society of London (Vol. 128, No. 3, pp. 431-454). Oxford, UK: Blackwell Publishing Ltd. https://doi.org/10.1111/j.1096-3642.1957.tb00336.x

South, A., 2017. rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0. https://CRAN.R-project.org/package=rnaturalearth

Spear, S.F., Groves, J.D., Williams, L.A. and Waits, L.P., 2015. Using environmental DNA methods to improve detectability in a hellbender (*Cryptobranchus alleganiensis*) monitoring program. Biological Conservation, 183, pp.38-45. https://doi.org/10.1016/j.biocon.2014.11.016

Spens, J., Evans, A.R., Halfmaerten, D., Knudsen, S.W., Sengupta, M.E., Mak, S.S., Sigsgaard, E.E. and Hellström, M., 2017. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution, 8(5), pp.635-645. https://doi.org/10.1111/2041-210X.12683

Sternecker, K., Denic, M., and Geist, J., 2014. Timing matters: species-specific interactions between spawning time, substrate quality, and recruitment success in three salmonid species. Ecology and Evolution, 4(13), pp.2749-2758. https://doi.org/10.1002/ece3.1128

Stewart, K.A., 2019. Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. Biodiversity and Conservation, 28(5), pp.983-1001. https://doi.org/10.1007/s10531-019-01709-8

Stoeckle, M.Y., Soboleva, L. and Charlop-Powers, Z., 2017a. Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. PloS one, 12(4), p.e0175186. https://doi.org/10.1371/journal.pone.0175186

Stoeckle, B. C., Beggel, S., Cerwenka, A. F., Motivans, E., Kuehn, R., Geist, J., 2017. A systematic approach to evaluate the influence of environmental conditions on eDNA detection success in aquatic ecosystems. PloS one, 12(12), e0189119. https://doi.org/10.1371/journal.pone.0189119

Stork, N.E., Samways, M.J. and Eeley, H.A., 1995. Inventorying and monitoring biodiversity. Trends in Ecology & Evolution, 11(1), pp.39-40.

Strayer, D.L. and Dudgeon, D., 2010. Freshwater biodiversity conservation: recent progress and future challenges. Journal of the North American Benthological Society, 29(1), pp.344-358. https://doi.org/10.1899/08-171.1

Strickler, K.M., Fremier, A.K. and Goldberg, C.S., 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation, 183, pp.85-92. https://doi.org/10.1016/j.biocon.2014.11.038

Sumner, M.D., 2016. ggpolypath: Polygons with Holes for the Grammar of Graphics. R package version 0.1.0. https://CRAN.R-project.org/package=ggpolypath

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. and Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular ecology, 21(8), pp.2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H. and Kawabata, Z.I., 2012. Estimation of fish biomass using environmental DNA. PloS one, 7(4), p.e35868. https://doi.org/10.1371/journal.pone.0035868

Takahara, T., Minamoto, T. and Doi, H., 2013. Using environmental DNA to estimate the distribution of an invasive fish species in ponds. PloS one, 8(2), p.e56584. https://doi.org/10.1371/journal.pone.0056584

Takahashi, S., Sakata, M.K., Minamoto, T. and Masuda, R., 2020. Comparing the efficiency of open and enclosed filtration systems in environmental DNA quantification for fish and jellyfish. PloS one, 15(4), p.e0231718. https://doi.org/10.1371/journal.pone.0231718

Takeuchi, A., Iijima, T., Kakuzen, W., Watanabe, S., Yamada, Y., Okamura, A., Horie, N., Mikawa, N., Miller, M.J., Kojima, T. and Tsukamoto, K., 2019. Release of eDNA by different life history stages and during spawning activities of laboratory-reared Japanese eels for interpretation of oceanic survey data. Scientific reports, 9(1), pp.1-9. https://doi.org/10.1038/s41598-019-42641-9

Telnes, T., and Saegrov, H., 2004. Reproductive strategies in two sympatric morphotypes of Arctic charr in Kalandsvatnet, west Norway. Journal of fish biology, 65(2), pp.574-579. https://doi.org/10.1111/j.0022-1112.2004.00432.x

Thalinger, B., Wolf, E., Traugott, M. and Wanzenböck, J., 2019. Monitoring spawning migrations of potamodromous fish species via eDNA. Scientific reports, 9(1), pp.1-11. https://doi.org/10.1038/s41598-019-51398-0

Thomsen, P.F., Kielgast, J.O.S., Iversen, L.L., Wiuf, C., Rasmussen, M., Gilbert, M.T.P., Orlando, L. and Willerslev, E., 2012a. Monitoring endangered freshwater biodiversity using environmental DNA. Molecular ecology, 21(11), pp.2565-2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Thomsen, P.F., Kielgast, J., Iversen, L.L., Møller, P.R., Rasmussen, M. and Willerslev, E., 2012b. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS one, 7(8), p.e41732. https://doi.org/10.1371/journal.pone.0041732

Thomsen, P.F. and Willerslev, E., 2015. Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. Biological conservation, 183, pp.4-18. https://doi.org/10.1016/j.biocon.2014.11.019.

Tillotson, M.D., Kelly, R.P., Duda, J.J., Hoy, M., Kralj, J. and Quinn, T.P., 2018. Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. Biological Conservation, 220, pp.1-11. https://doi.org/10.1016/j.biocon.2018.01.030

Tomlinson, M.L. and Perrow, M.R., 2003. Ecology of the Bullhead: *Cottus gobio*. Life in UK Rivers.

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T. and Yamanaka, H., 2017. Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. PLoS One, 12(4), p.e0176608. https://doi.org/10.1371/journal.pone.0176608

Tsuji, S. and Shibata, N., 2020. Identifying spawning activity in aquatic species based on environmental DNA spikes. BioRxiv.

Turner, C.R., Miller, D.J., Coyne, K.J. and Corush, J., 2014. Improved methods for capture, extraction, and quantitative assay of environmental DNA from Asian

bigheaded carp (*Hypophthalmichthys* spp.). PloS one, 9(12), p.e114329.
https://doi.org/10.1371/journal.pone.0114329

Turner, C.R., Uy, K.L. and Everhart, R.C., 2015. Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation, 183, pp.93-102. https://doi.org/10.1016/j.biocon.2014.11.017

UK Lakes Portal. Environmental Information Platform. Available at:
https://eip.ceh.ac.uk/apps/lakes/. Retrieved August 2020.

UKTAG, 2006. UK Environmental standards and conditions. UK Technical Advisory Group on the Water Framework Directive. January 2006.

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., Yamanaka, H., Minamoto, T. and Kondoh, M., 2018. Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. Metabarcoding and Metagenomics, 2, p.e23297. https://doi.org/10.3897/mbmg.2.23297

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F. and Gaboriaud, C., 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular ecology, 25(4), pp.929-942. https://doi.org/10.1111/mec.13428

Veech, J.A., 2013. A probabilistic model for analysing species co-occurrence. Global Ecology and Biogeography, 22(2), pp.252-260. https://doi.org/10.1111/j.1466-8238.2012.00789.x

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. Springer.

Willby, N., Pitt, J.-A., Phillips, G., 2012. The ecological classification of UK lakes using aquatic macrophytes. Report – SC010080/R2. ISBN: 978-1-84911-286-4.

Willby, N., Law, A., Bull, C., Hänfling, B., Lawson Handley, L., Winfield, J.I., 2019. A tool for classifying the ecological status of lake fish in Britain based on eDNA metabarcoding. Report to Scottish Environement Protection Agency.

Williams, M. A., O'Grady, J., Ball, B., Carlsson, J., de Eyto, E., McGinnity, P., Jennings, E. Regan, F., Parle-McDermott, A., 2019. The application of CRISPR-Cas for single species identification from environmental DNA. Molecular Ecology Resources, 19(5), pp.1106-1114. https://doi.org/10.1111/1755-0998.13045

Winfield, I.J., 2002. Monitoring lake fish communities for the Water Framework Directive: a UK perspective. TemaNord, 566, pp.69-72.

Winfield, I.J. and Durie, N.C., 2004. Fish introductions and their management in the English Lake District. Fisheries Management and Ecology, 11(3-4), pp.195-201. https://doi.org/10.1111/j.1365-2400.2004.00399.x

Winfield, I.J., Fletcher, J.M. and James, J.B., 2004, January. Conservation ecology of the vendace (*Coregonus albula*) in Bassenthwaite Lake and Derwent Water, UK. In Annales Zoologici Fennici (pp. 155-164). Finnish Zoological and Botanical Publishing Board. http://www.jstor.org/stable/23736198

Winfield, I. J.; Fletcher, J. M.; James, J. B.; Bayliss, B. D., 2006. Fisheries on the edge in Cumbria UK: where salmonids, cyprinids and climate change collide. In: Fisheries on the edge. Proceedings of the Institute of Fisheries Management Conference, 2005. Nottingham, Insititute of Fisheries Management, pp.125-136.

Winfield, I.J., Fletcher, J.M. and James, J.B., 2007. Seasonal variability in the abundance of Arctic charr (*Salvelinus alpinus* (L.)) recorded using hydroacoustics in Windermere, UK and its implications for survey design. Ecology of Freshwater Fish, 16(1), pp.64-69. https://doi.org/10.1111/j.1600-0633.2006.00170.x

Winfield, I.J., Fletcher, J.M. and James, J.B., 2008a. The Arctic charr (*Salvelinus alpinus*) populations of Windermere, UK: population trends associated with eutrophication, climate change and increased abundance of roach (*Rutilus rutilus*). Environmental Biology of Fishes, 83(1), pp.25-35. https://doi.org/10.1007/s10641-007-9235-4

Winfield, I.J., Fletcher, J.M. and Lyle, A.A., 2008b. Assessment of the vendace refuge population of Loch Skeen. Final Report.

Winfield, I.J., Fletcher, J.M., James, J.B. and Bean, C.W., 2009. Assessment of fish populations in still waters using hydroacoustics and survey gill netting: experiences with Arctic charr (*Salvelinus alpinus*) in the UK. Fisheries Research, 96(1), pp.30-38. DOI: 10.1016/j.fishres.2008.09.013

Winfield, I.J., Hateley, J., Fletcher, J.M., James, J.B., Bean, C.W. and Clabburn, P., 2010. Population trends of Arctic charr (*Salvelinus alpinus*) in the UK: assessing the evidence for a widespread decline in response to climate change. Hydrobiologia, 650(1), pp.55-65. https://doi.org/10.1007/s10750-009-0078-1

Winfield, I.J., Fletcher, J.M. and James, J.B., 2011. Rare fish monitoring. Final Report.

Winfield, I.J., Adams, C.E., Bean, C.W., Durie, N.C., Fletcher, J.M., Gowans, A.R., Harrod, C., James, J.B., Lyle, A.A., Maitland, P.S. and Thompson, C., 2012. Conservation of the vendace (*Coregonus albula*), the UK's rarest freshwater fish. Advances in Limnology, 63, pp.547-559.

Winfield, I.J., Bean, C.W., Gorst, J., Gowans, A.R., Robinson, M. and Thomas, R., 2013. Assessment and conservation of whitefish (*Coregonus lavaretus*) in the UK. Advances in Limnology, 64, pp.305-321. DOI: 10.1127/1612-166X/2013/0064-0023

Winfield, I.J., van Rijn, J. and Valley, R.D., 2015. Hydroacoustic quantification and assessment of spawning grounds of a lake salmonid in a eutrophicated water body. Ecological Informatics, 30, pp.235-240. https://doi.org/10.1016/j.ecoinf.2015.05.009

Winfield, I.J., Fletcher, J.M. and James, J.B., 2017. The 'reappearance' of vendace (*Coregonus albula*) in the face of multiple stressors in Bassenthwaite Lake, UK. Fundamental and Applied Limnology, 189(3), pp.227-233. https://doi.org/10.1127/fal/2016/0799

Winfield, I.J., Berry, R. and Iddon, H., 2019. The cultural importance and international recognition of the Arctic charr *Salvelinus alpinus* populations of Windermere, UK. Hydrobiologia, 840(1), pp.11-19. https://doi.org/10.1007/s10750-018-3814-6

Wood, S.A., Pochon, X., Laroche, O., von Ammon, U., Adamson, J. and Zaiko, A., 2019. A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and

metabarcoding for species-specific detection in environmental DNA. Molecular ecology resources, 19(6), pp.1407-1419.https://doi.org/10.1111/1755-0998.13055

Yamahara, K. M., Preston, C. M., Birch, J., Walz, K., Marin III, R., Jensen, S., Pargett, D., Roman, B., Ussler III, W., Zhang, Y., Ryan, J., Hobson, B., Kieft, B., Raanan, B., Goodwin, K.D., Chavez, F.P., Scholin, C., 2019. *In situ* autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. Frontiers in Marine Science, 6, 373. https://doi.org/10.3389/fmars.2019.00373

Yatsuyanagi, T., Ishida, R., Sakata, M.K., Kanbe, T., Mizumoto, H., Kobayashi, Y., Kamada, S., Namba, S., Nii, H., Minamoto, T. and Araki, H., 2020. Environmental DNA monitoring for short-term reproductive migration of endemic anadromous species, Shishamo smelt (*Spirinchus lanceolatus*). Environmental DNA, 2(2), pp.130-139. https://doi.org/10.1002/edn3.50

Yusishen, M.E., Eichorn, F.C., Anderson, W.G. and Docker, M.F., 2020. Development of quantitative PCR assays for the detection and quantification of lake sturgeon (*Acipenser fulvescens*) environmental DNA. Conservation Genetics Resources, 12(1), pp.17-19. https://doi.org/10.1007/s12686-018-1054-8

Zhang, S., Lu, Q., Wang, Y., Wang, X., Zhao, J. and Yao, M., 2020. Assessment of fish communities using environmental DNA: Effect of spatial sampling design in lentic systems of different sizes. Molecular Ecology Resources, 20(1), pp.242-255. https://doi.org/10.1111/1755-0998.13105

Appendix 1: Additional information for Chapter 2

Supplementary methods

DNA extraction from Sterivex filters preserved with buffers

The following protocol describes separate DNA extractions from Sterivex units (hereafter STX-f) and preservation buffers (hereafter STX-b; ethanol, Longmire's and RNAlater), however, DNA extracted from filer unit and buffer of the same sample can be combined (see protocol for details).

This protocol is modified from Sellers et al., 2018 and details of reagents used here and their preparation can be found at: https://doi.org/10.17504/protocols.io.qn9dvh6

The entire protocol is scalable based upon transferred supernatant volumes. Where relevant, scalable volumes are shown in brackets next to reagent volumes used (i.e. 2X volume).

Important: incubate *Binding Solution* at 55°C until required. If any precipitate is present mix gently until dissolved. Use solutions while still warm.

LYSIS

1. Samples preparation:

Carefully wipe the outer surfaces of all filter units with 10% bleach using clean tissue paper. Dry and wipe with 70% ethanol using tissue paper.

Transfer the buffer from the filter unit and into a 1.5 mL sterile tube using a 5 mL Luer Lock syringe. Be careful not to apply too much pressure.

2. Centrifugation step:

2a *STX-f*

Remove the inlet and the outlet caps. Place the filters vertically with the 'inlet end' facing down in a 50 mL tube and clean laboratory tissue paper underneath. Centrifuge

at 4,000 x g for 1 min at room temperature. After drying follow the exact procedure of step 3a.

2b *STX-b*

Spin at 6,000 x g for 30-45 mins in a micro-centrifuge at room temperature. Discard liquid. Let pellet dry. Go to step 3b.

3. Addition of lysis buffer:

3a *STX-f*

Keep the outlet end closed with the outlet cap. Carefully add 780 μL of *Lysis Solution*, 60 μL of *20% SDS* and 60 μL of *Proteinase K* to the filter by using a 1,000 μL pipet and sterile filter tips (<u>*note:*</u> prepare lysis master mix in advance). Pipet the solution between the outside of the filter and the capsule walls. Close with an inlet cap, seal with parafilm. Handshake vigorously for a few seconds. Proceed to step 4.

3b *STX-b*

Dissolve the dried pellet by using 260 μL of *Lysis Solution*, 20 μL of *20% SDS* and 20 μL of *Proteinase K*. Close tube and seal with parafilm. Vortex for 15 s and proceed to step 5.

4. Incubate: while rotating at 56°C overnight (<u>*note:*</u> STX-f should lay horizontally into a rosetta type apparatus rotating overnight; STX-b can be placed into a rack over a heated rocking platform).

5. Handshake STX-f vigorously 5 times and vortex STX-b for 15 s.

6. Transfer:

6a *STX-f*

Remove the lysate from the inlet end of the capsule by using a Luer Lock syringe. Transfer the lysate to a 1.5 μL tube. Vortex for a few seconds. Spin down the excess debris at 4,000 x g for 5 seconds and transfer the clean lysate to a 1.5 mL tube. Measure the volume.

6b *STX-b*

Spin down the excess debris at 4,000 x g for 5 seconds and transfer the clean lysate to a 1.5 mL tube. Measure the volume.

Note: lysate from STX-f and STX-b of the same sample can be combined here.

INHIBITOR REMOVAL

The following steps are based on 500 - 650 µL of transferred lysate.

1. Add 300 µL (0.6X volume) of *Water Flocculant Solution*, vortex briefly and incubate at 4°C for 30 mins or on ice for a minimum of 10 mins.
2. Centrifuge at 10,000 x g for 1 min at room temperature.
3. Without disturbing the pellet, transfer the supernatant to a 2 mL tube.

SILICA BINDING

The following steps are based on 600 - 700 µL of transferred supernatant.

1. Add 1,200 µL (2X volume) of *Binding Solution*, vortex briefly to mix.
2. Transfer 650 µL of the mixture to a spin column.
3. Centrifuge at ≥ 10,000 x g for 1 min at room temperature, discard the flow-through.
4. Repeat steps 2 to 3 until all the mixture has passed through the spin column.

WASH

1. Add 500 µL of *Wash Solution* to the spin column.
2. Centrifuge at 10,000 x g for 1 min at room temperature, discard the flow-through.
3. Repeat steps 1 and 2 a second time.
4. Centrifuge at 10,000 x g for 2 mins at room temperature, replace collection tube with a fresh 1.5 mL tube.

ELUTION

1. Add 100 μL of 70°C *Elution Buffer* directly to the spin column membrane and incubate for 10 mins at room temperature.

2. Centrifuge at 10,000 x g for 1 min at room temperature.

3. DNA is now in the 1.5 mL tube.

Optional: For increased DNA yield repeat steps 1 and 2 a further time.

PCR Clean-up: double-size selection magnetic beads protocol

The protocol from Bronner *et al.* (2013) will be used with minor modifications.
A ratio of 0.9x and 0.15x Mag-Bind® RXNPure Plus magnetic beads to 100 μl PCR product will be used. The 0.9x ratio binds any product from 300 bp upwards whilst the second 0.15x ratio will bind larger secondary products at ~1000 bp plus. Therefore, 90 μl and 15 μl of magnetic beads will be required for 100 μl PCR product. After clean-up, use gel electrophoresis to check cleaned sub-libraries against their non-cleaned PCR products to ensure clean-up has worked.

Note: for second PCR clean-up the protocol described below is used, but the first bead ratio need to be adjusted to 0.7x.

1. Place 1.5 ml microcentrifuge tubes on tube rack.

2. Allow Mag-Bind® RXNPure Plus beads to come to room temperature before use. Vortex the Mag-Bind® RXNPure Plus aliquot to resuspend beads. Add 0.9x (90 μl) Mag-Bind® RXNPure Plus to a microcentrifuge tube for each sub-library.

3. Add 100 μl of PCR product for each sub-library. Pipette up and down 5-10 times and vortex for 30 seconds to mix well. Briefly spin down for 2 seconds to collect residue on side of tube.

4. Incubate DNA/bead mixture at room temperature for 5 minutes.

5. Place tubes on a magnetic stand to bind the Mag-Bind® RXNPure Plus beads to side of the tube. Wait for solution to clear of all Mag-Bind® RXNPure Plus beads (~1 min).

6. Carefully transfer the supernatant to a new tube and discard the tube containing the beads (DO NOT discard the supernatant).

7. Add 0.15x (15 μl) Mag-Bind® RXNPure Plus beads to the supernatant, vortex to mix well, and briefly spin down for 2 seconds to collect residue on side of tube.

8. Incubate DNA/bead mixture for 5 minutes at room temperature.

9. Put tubes on a magnetic stand and wait for the solution to clear (~1 min).

10. Remove the cleared supernatant from the tube (DO NOT disturb the pellet on the side).

11. Add 200 µl 80% ethanol to each tube without disturbing the pellet. Wait 30 seconds and discard the ethanol (DO NOT disturb the pellet on the side).

12. Repeat Step 11.

13. Remove residual ethanol with 10/20 µl pipette tips.

14. Leave tubes with lids open on the magnetic stand for 10-15 minutes to dry the Mag-Bind® RXNPure Plus beads. Remove residual liquid with 10/20 µl pipette tips.

15. Remove tubes from the magnetic stand.

16. Add 25 µl Elution Buffer to each tube. Pipette up and down 20 times or vortex for 30 seconds to mix well.

17. Incubate at room temperature for 5 minutes.

18. Put tubes back on the magnetic stand and wait for the solution to clear (~2-3 min).

19. Transfer the cleared supernatant (23.5 µl) to new 1.5mL microcentrifuge tubes.

20. Store purified DNA at 2-8°C if storage is only for a few days. For long-term storage, samples should be kept at -20°C.

Supplementary table and figure

Table S1. List of fish species and their abundance/biomass in New Lake (NL) and Middle Lake (ML).

| Pond	Common name	Scientific name	Abundance (number of individuals)	Total biomass (kg)
NL	Common carp	*Cyprinus carpio*	293	852.8
NL	Common carp x crucian carp	*C. carpio x Carassius carassius*	17	16
NL/ML	Bream	*Abramis brama*	382	240.6
NL/ML	Barbel	*Barbus barbus*	3	3.1
NL/ML	Crucian carp	*C. carassius*	378	123.7
NL/ML	Common carp x crucian carp	*C. carpio x C. carassius*	7	2.9
NL/ML	European perch	*Perca fluviatilis*	23	10.57
NL/ML	Roach x bream	*Rutilus rutilus x A. brama*	60	31.4
NL/ML	Roach	*R. rutilus*	184	46.9
NL/ML	Rudd	*Scardinius erythrophthalmus*	7	1.5
NL/ML	Chub	*Squalius cephalus*	1	0.7
NL/ML	Tench	*Tinca tinca*	203	173.5
NL	Crucian carp	*C. carassius*	1	0.2
NL	Grass carp	*Ctenopharyngodon idella*	4	17.9
NL	Common carp	*C. carpio*	190	862.4
NL	Common carp x crucian carp	*C. carpio x C. carassius*	4	1.8
NL	European perch	*Perca fluviatilis*	5	1.2
NL	Bream	*A. brama*	151	42.7
NL	Sterlet	*Acipenser* spp.	1	1
NL	Roach x bream	*R. rutilus x A. brama*	5	1.8
NL	Roach	*R. rutilus*	58	10.7
NL	Wels catfish	*Silurus glanis*	23	251.95
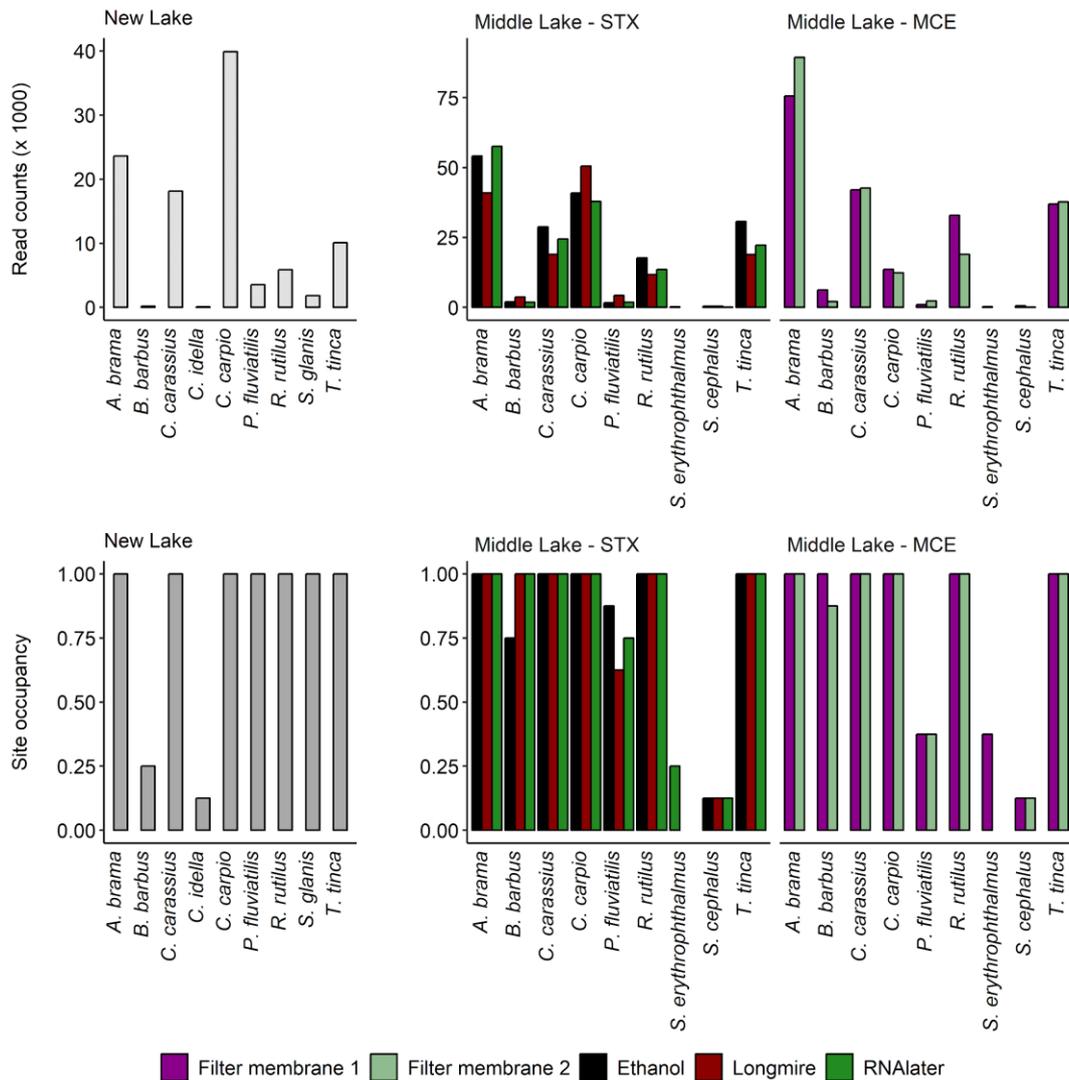
Figure S1. Fish species read counts (top) and site occupancy (bottom) barplots for New Lake, Middle Lake Sterivex sampling (Middle Lake-STX) and Middle Lake MCE open filters sampling (Middle Lake-MCE). *Note* mixed samples (pooled aliquots of individual field samples) were not included in site occupancy calculations.

Appendix 2: Additional information for Chapter 3

Supplementary table and figures

Table S1. Coordinates (latitude and longitude) in decimal degrees of the eDNA collection sites at Windermere north basin for the locations of North Thompson Holme (SNTH), Red Nab (SRN), and offshore transect (OF).

Sampling sites	Latitude (decimal degrees)	Longitude (decimal degrees)
SNTH 1	54.366660°	-2.939000°
SNTH 2	54.366581°	-2.937518°
SNTH 3	54.366890°	-2.936380°
OF 1	54.368870°	-2.936110°
OF 2	54.370850°	-2.935830°
OF 3	54.372830°	-2.935270°
OF 4	54.374820°	-2.935000°
OF 5	54.376800°	-2.934440°
SRN 1	54.364900°	-2.943600°
SRN 2	54.369800°	-2.943100°
SRN 3	54.372462°	-2.943337°
SRN 4	54.375735°	-2.943932°
SRN 5	54.378554°	-2.944028°
SRN 6	54.381511°	-2.943968°
SRN 7	54.384358°	-2.945170°
SRN 8	54.386462°	-2.947511°
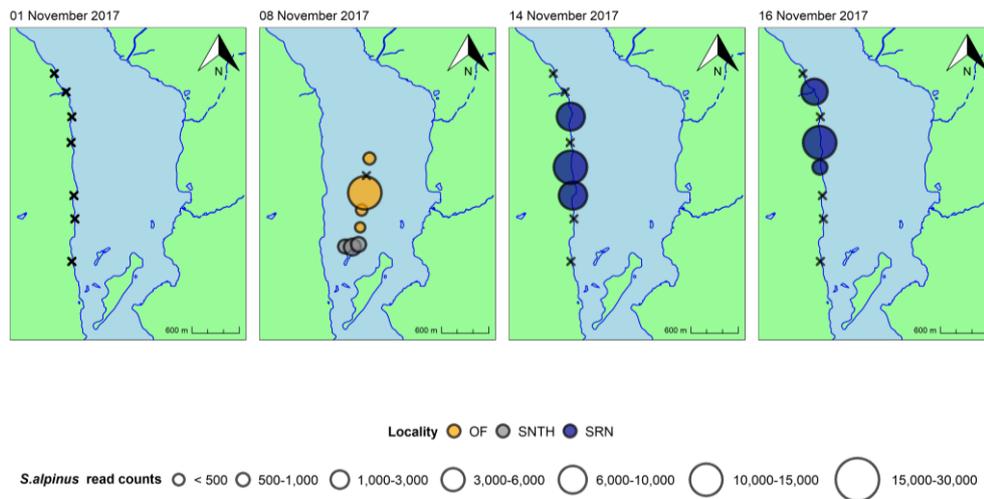
Figure S1. Spatio-temporal variation of Arctic charr eDNA signal in Windermere's north basin. Bubble size is proportional to read counts assigned to Arctic charr, whereas black crosses indicate sites where the species was not detected. Maps were created using shape files downloaded from EDINA Digimap® Ordinance Survey service (http://edina.ac.uk/digimap).
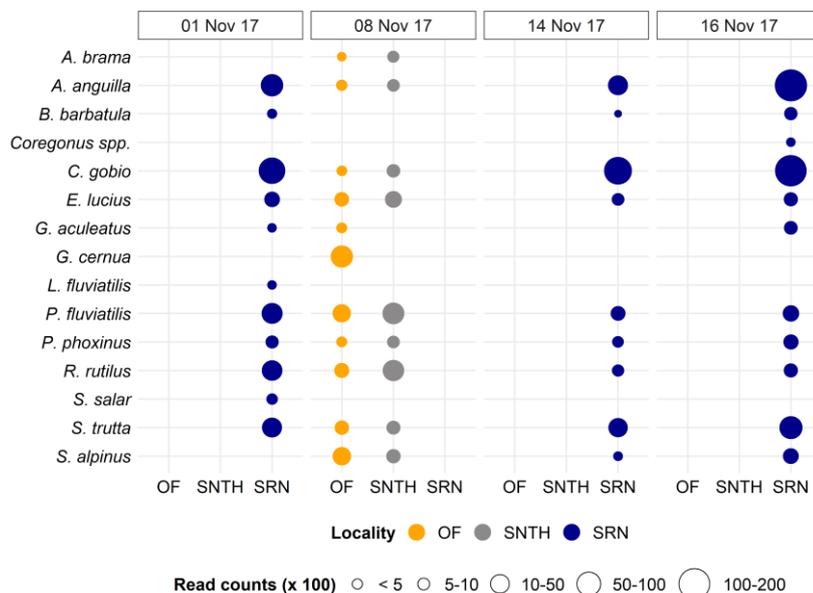
Figure S2. Community plot of eDNA metabarcoding detections and relative abundance of fish in Lake Windermere. Bubble size show average read counts of fish species detected over additional eDNA sampling dates in Windermere's north basin. SRN includes average read counts of species detected at the shoreline sites of Red Nab (*n* = 8), SNTH includes sites at the shore of North Thompson island (*n* = 3) and OF includes offshore sites between 20 and 40 m depth (*n* = 5).

172

Appendix 3: Additional information for Chapter 4

Supplementary tables

Table S1. Bayesian estimates of Arctic charr detection probability in metabarcoding PCR replicates (p) given that the species is present at the sites (ψ) and in samples (ϑ).

Site (ψ)	Sample (ϑ)	Posterior median (p)	95% CI	Vertebrate DNA (ng/µL)	Total DNA (ng/µL)
OFF	1	0,62	0,50 - 0,73	0,000613	10,3
OFF	2	0,61	0,53 - 0,69	0,001181	7,72
OFF	3	0,62	0,50 - 0,73	0,055214	10,4
OFF	4	0,61	0,53 - 0,68	0,001357	7,15
OFF	5	0,65	0,55 - 0,74	1,909915	4,4
OFF	6	0,62	0,53 - 0,70	0,503837	5,27
OFF	7	0,63	0,56 - 0,69	0,829924	6,59
OFF	8	0,63	0,57 - 0,70	0,897546	7,98
OFF	9	0,62	0,55 - 0,69	0,444075	6,55
SNTH	1	0,61	0,53 - 0,68	0,001058	6,77
SNTH	2	0,62	0,50 - 0,73	0,001252	10,27
SNTH	3	0,61	0,53 - 0,68	0,000712	6,91
SNTH	4	0,63	0,57 - 0,69	0,79293	7,51
SNTH	5	0,65	0,58 - 0,71	1,358538	8,95
SNTH	6	0,64	0,57 - 0,71	1,419588	6,11
SRN	1	0,75	0,65 - 0,85	5,627803	11,57
SRN	2	0,71	0,64 - 0,78	4,111025	8,25
SRN	3	0,72	0,64 - 0,79	4,574705	7,95
SRN	4	0,69	0,62 - 0,75	3,334065	7,09
SRN	5	0,72	0,58 - 0,83	4,723269	4,34
SRN	6	0,68	0,55 - 0,80	3,380479	3,76
SRN	7	0,85	0,66 - 0,96	10,76755	16,4
SRN	8	0,77	0,66 - 0,86	6,628753	10,03

Table S2. Bayesian estimates of Arctic charr detection probability in qPCR replicates (p) given that the species is present at the sites (ψ) and in samples (ϑ).

| Site (ψ)	Sample (ϑ)	Posterior median (p)	95% CI	Vertebrate DNA (ng/µL)	Total DNA (ng/µL)
OFF	1	0,58	0,47 - 0,69	0,000613	10,3
OFF	2	0,55	0,47 - 0,63	0,001181	7,72
OFF	3	0,59	0,47 - 0,70	0,055214	10,4
OFF	4	0,54	0,46 - 0,62	0,001357	7,15
OFF	5	0,55	0,44 - 0,65	1,909915	4,4
OFF	6	0,53	0,44 - 0,61	0,503837	5,27
OFF	7	0,55	0,48 - 0,62	0,829924	6,59
OFF	8	0,57	0,50 - 0,64	0,897546	7,98
OFF	9	0,54	0,47 - 0,61	0,444075	6,55
SNTH	1	0,54	0,46 - 0,61	0,001058	6,77
SNTH	2	0,58	0,47 - 0,69	0,001252	10,27
SNTH	3	0,54	0,46 - 0,61	0,000712	6,91
SNTH	4	0,56	0,50 - 0,63	0,79293	7,51
SNTH	5	0,60	0,53 - 0,66	1,358538	8,95
SNTH	6	0,56	0,49 - 0,63	1,419588	6,11
SRN	1	0,72	0,61 - 0,81	5,627803	11,57
SRN	2	0,65	0,57 - 0,72	4,111025	8,25
SRN	3	0,65	0,57 - 0,73	4,574705	7,95
SRN	4	0,61	0,54 - 0,68	3,334065	7,09
SRN	5	0,61	0,47 - 0,74	4,723269	4,34
SRN	6	0,57	0,44 - 0,70	3,380479	3,76
SRN	7	0,85	0,67 - 0,95	10,76755	16,4
SRN	8	0,72	0,61 - 0,82	6,628753	10,03

Supplementary figures
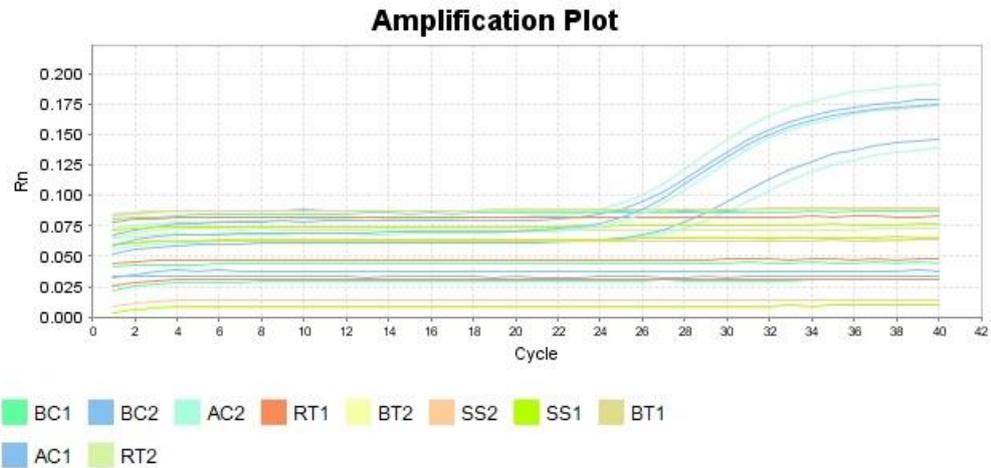
Figure S1. Amplification plot of Arctic charr qPCR assay showing specificity for the amplification of the target species (AC1 and AC2). Brook trout (*Salvelinus fontinalis*; BC1 and BC2), rainbow trout (*Onchorhynchus mykiss*; RT1 and RT2), brown trout (*Salmo trutta*; BT1 and BT2), and Atlantic salmon (*Salmo salar*; SS1 and SS2) did not show amplification with this assay. To test the specificity of the Arctic charr qPCR assay, genomic DNA of the five species tested was standardised to 1 ng/ µL.
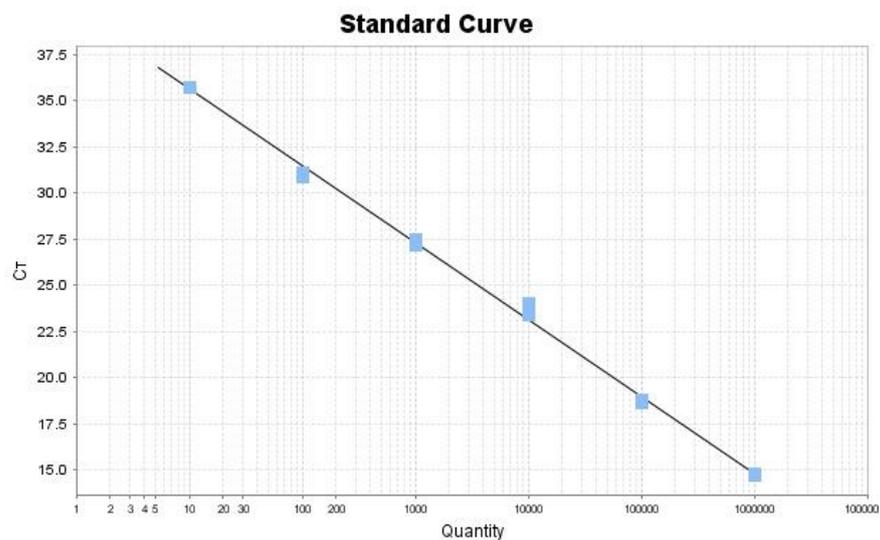
Figure S2. Standard curve of the Arctic charr qPCR assay tested for sensitivity. The determination of the assay sensitivity was determined using 10-fold dilutions of a 400 bp gBlocks® Gene Fragment ranging from 5×10^6 to 5 copies/µL. The limit of detection (the concentration at which no Arctic charr DNA amplifies) does not appear in the plot, whereas the limit of quantification (the concentration at which Arctic charr DNA amplifies inconsistently across replicates) was estimated to 5×10 copies/µL.
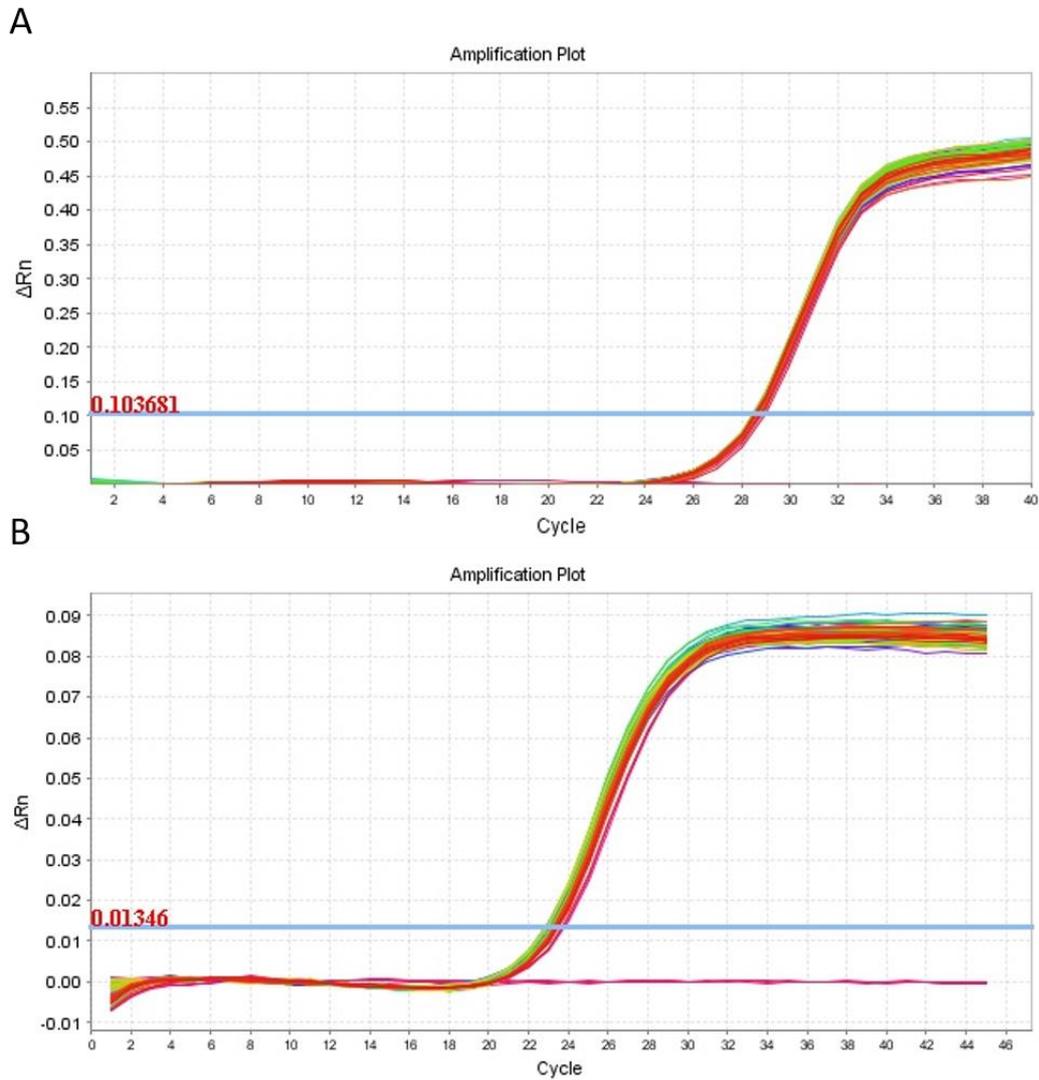
Figure S3. Amplification plots of qPCR inhibition test using exogenous internal positive control (EXO-IPC; **A**) and spiked Arctic charr genomic DNA (**B**). Light blue lines indicate the instrument thresholds for detections of PCR products during the amplification cycles. In both figures samples not showing amplification are negative controls (no DNA template added).
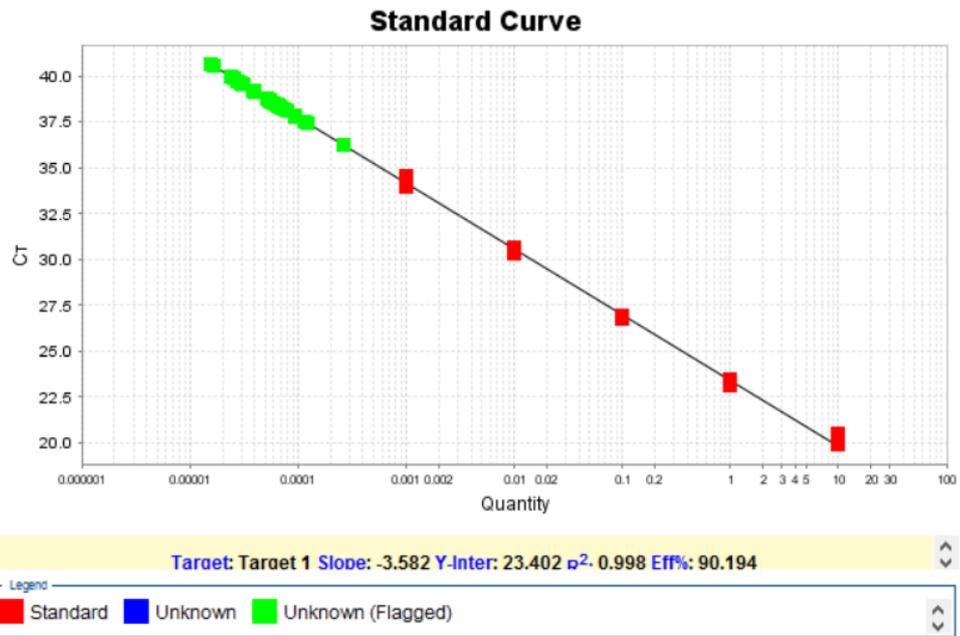
Figure S4. Example of qPCR standard curve showing eDNA samples (in green) used in the study and amplifying below the limit of quantification defined by the standards (in red).

Appendix 4: Additional information for Chapter 5

Supplementary tables

Table S1. Abundance scores assigned to Arctic charr and coregonids (vendace or European whitefish) at each site surveyed based on historical data and site occupancy from eDNA metabarcoding. Number of shore and offshore (when applicable) water samples are reported.

| Location	Shoreline samples	Offshore samples	Arctic charr - Historical data score / eDNA score	Coregonids – Historical data score / eDNA score
Bassenthwaite Lake	20	NA	0/0	1/1
Brothers Water	20	NA	0/0	4/1
Derwent Water	20	NA	0/0	4/1
Loch Awe	20	20	4/2	0/1
Loch Earn	22	18	5/5	1/3
Loch Eck	20	NA	1/1	5/5
Loch Lomond	30	40	0/1	4/5
Loch Skeen	10	10	0/2	5/5
Ullswater	20	NA	0/0	4/5
Buttermere	20	NA	2/4	0/0
Coniston Water	20	NA	4/3	0/0
Crummock Water	20	NA	4/5	0/0
Ennerdale Water	19	NA	4/4	0/0
Kielder Water	11	NA	1/1	0/0
Loch Ashie	20	NA	5/1	0/0
Loch Calder	20	NA	2/1	0/0
Loch Duntelchaig	20	NA	3/5	0/0
Loch Insh	20	NA	5/4	0/0
Loch Lee	20	NA	5/5	0/0
Loch Lubhair	20	NA	0/1	0/0
Loch Lubnaig	10	20	4/3	0/0

Loch Osgaig	10	10	2/1	0/0
Loch Rannoch	20	NA	4/3	0/0
Loch Tarff	20	NA	4/4	0/0
Loch Tay	20	NA	4/1	0/0
Loch Voil	20	NA	4/3	0/0
Loch Naver	20	NA	3/4	0/0
Loch Stack	20	NA	5/4	0/0
Wastwater	20	NA	2/2	0/0
Windermere	40	50	3/2	0/1
Loch Doon	20	NA	3/1	0/0
Loch Grannoch	20	NA	0/0	0/0
Oss Mere	10	10	0/0	0/0
Loch Achray	20	NA	0/0	0/0
Loch Eilt	20	NA	0/0	0/0
Loch Katrine	19	18	0/0	0/0
Loweswater	20	NA	0/0	0/0
Loch Affric	20	NA	0/0	0/0
Grimwith Reservoir	9	8	1/0	0/0
Sprinkling Tarn	10	NA	0/5	1/0
Llyn Padarn	10	10	4/2	0/0
Llyn Cwellyn	10	10	5/5	0/0
Betley Mere	10	10	0/0	0/0
Blelham Tarn	17	NA	0/0	0/0
Aqualate Mere	20	NA	0/0	0/0
Trinity Broads	20	NA	0/0	0/0
Loch Ussie	20	NA	0/0	0/0
Bomere Pool	20	NA	0/0	0/0
Broomlee Lough	20	NA	0/0	0/0
Burnmoor Tarn	20	NA	0/0	0/0
Chapel Mere	10	10	0/0	0/0
Crag Lough	20	NA	0/0	0/0
Crose Mere	20	NA	0/0	0/0
Dozmary Pool	20	NA	0/0	0/0
Elterwater	20	NA	0/0	0/0

Esthwaite Water	20	NA	0/0	0/0
Fenemere	10	10	0/0	0/0
Fleet Pond	20	NA	0/0	0/0
Grasmere	20	NA	0/0	0/0
Greatstone Lake	20	NA	0/0	0/0
Greenlee Lough	20	NA	0/0	0/0
Groby Pool	20	NA	0/0	0/0
Hawes Water	20	NA	0/0	0/0
Hornsea Mere	20	NA	0/0	0/0
Kenfig Pool	12	NA	0/0	0/0
Little Sea	20	NA	0/0	0/0
Llan Bwch-llyn Lake	10	NA	0/0	0/0
Llangorse Lake	10	10	0/0	0/0
Llyn Ogwen	20	NA	0/0	0/0
Llyn Penrhyn	10	NA	0/0	0/0
Llyn Traffwll	20	NA	0/0	0/0
Loch Davan	20	NA	0/0	0/0
Loch Dee	20	NA	0/0	0/0
Loch Doilet	20	NA	0/0	0/0
Loch Druidibeag	20	NA	0/0	0/0
Loch Fitty	20	NA	0/0	0/0
Loch Gorm	20	NA	0/0	0/0
Loch Ken	20	NA	0/0	0/0
Loch Kinord	20	NA	0/0	0/0
Loch Leven	10	20	0/0	0/0
Loch Morlich	20	NA	0/0	0/0
Loch of Harray	20	NA	0/0	0/0
Loch of Strathbeg	20	NA	0/0	0/0
Loch of Swannay	20	NA	0/0	0/0
Loch Ore	20	NA	0/0	0/0
Loch Scarmclate	20	NA	0/0	0/0
Loch Urghag	20	NA	0/0	0/0
Lochindorb	20	NA	0/0	0/0
Maer Pool	10	10	0/0	0/0

Malham Tarn	20	NA	0/0	0/0
Over Water	20	NA	0/0	0/0
Scoat Tarn	20	NA	0/0	0/0
Semer Water	20	NA	0/0	0/0
Slapton Ley	20	NA	0/0	0/0
St Mary's Loch	20	NA	0/0	0/0
Tatton Mere	20	NA	0/0	0/0
The Loe	20	NA	0/0	0/0
The Mere	20	NA	0/0	0/0
Tindale Tarn	20	NA	0/0	0/0
Upton Broad	20	NA	0/0	0/0
Watch Lane Flash	10	10	0/0	0/0

Table S2. Pairwise probability table of significant species associations where: *obs_cooccur* is the observed number of sites with both species; *exp_cooccur* is the expected number of sites with both species; *p_lt* is the probability of species co-occurrence at a frequency less than the observed number of co-occurrence sites if the two species were distributed randomly of one another; *p_gt* is the probability of species co-occurrence at a frequency greater than the observed frequency. The values of *p_lt* and *p_gt* can be considered as *p*-values indicating significance levels of co-occurring patterns.

|	obs_cooccur	exp_cooccur	p_lt	p_gt
Arctic charr				
Rudd	0	6.2	0.00021	1.00000
Tench	2	6.8	0.00864	0.99876
Coregonids	6	3	0.99313	0.03690
Common bleak	0	3.6	0.01074	1.00000
Common carp	1	5	0.01331	0.99870
Pike	10	16.3	0.00526	0.99867
European river lamprey	18	10.4	0.99988	0.00066
European perch	16	20.5	0.03213	0.98946
Eurasian minnow	26	17.5	0.99999	0.00012
Roach	6	13.7	0.00067	0.99988
Atlantic salmon	25	12.5	1.00000	0.00000
Brown trout	30	23.8	1.00000	0.00021
Coregonids				
Stone loach	8	2.1	1.00000	0.00003
Three-spined stickleback	10	6.3	1.00000	0.00788
European river lamprey	9	3.5	0.99999	0.00025
Rainbow trout	6	2.6	0.99745	0.01712
Eurasian minnow	10	5.8	1.00000	0.00327
Nine-spined stickleback	4	1.3	0.99716	0.02301
Atlantic salmon	9	4.2	0.99992	0.00145
Arctic charr	6	3	0.99313	0.03690

Supplementary figures
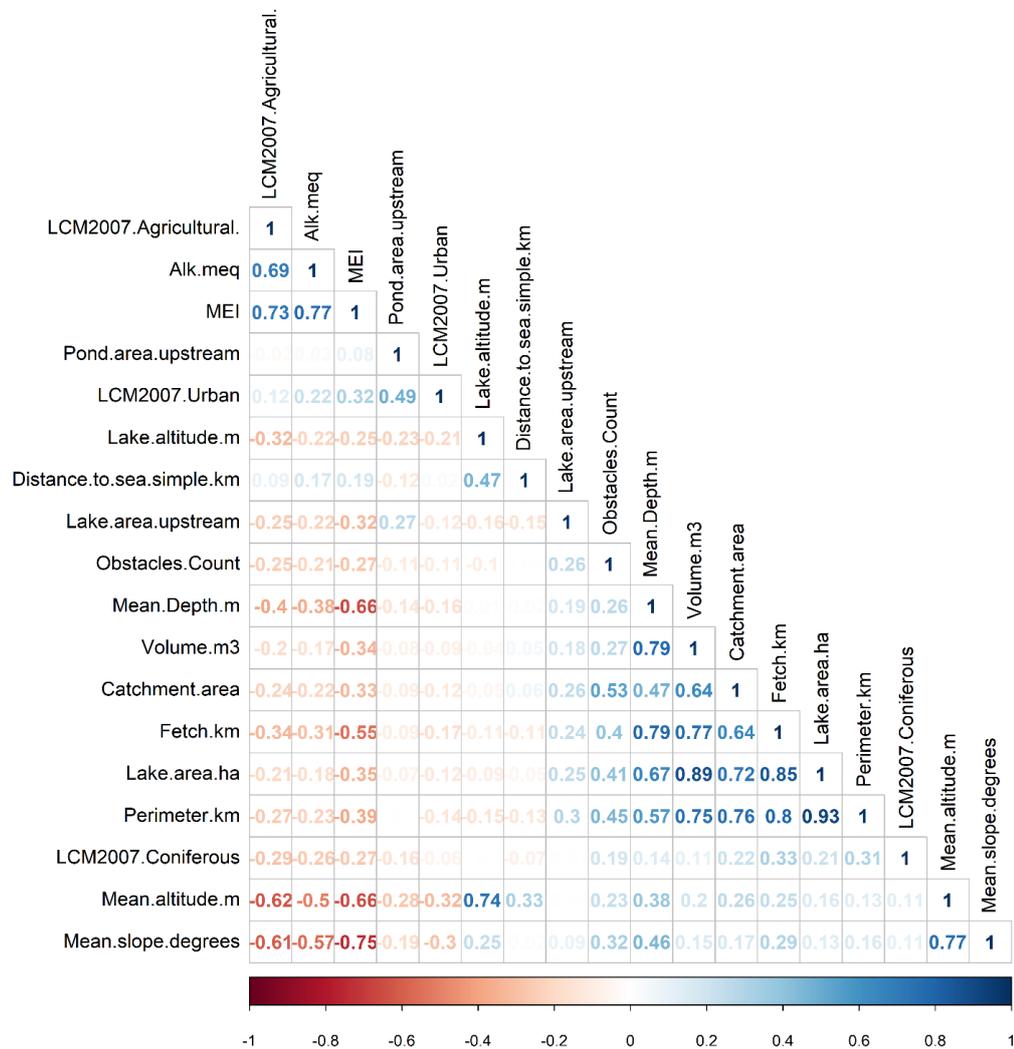
Figure S1. Correlation plot displaying Pearson's correlation coefficients between independent variables selected for GLM analysis. Positive correlations are displayed in blue and negative correlations in red, and colour intensity is proportional to the correlation coefficients. Only variables with coefficients lower than 0.35 were kept to build the full model.
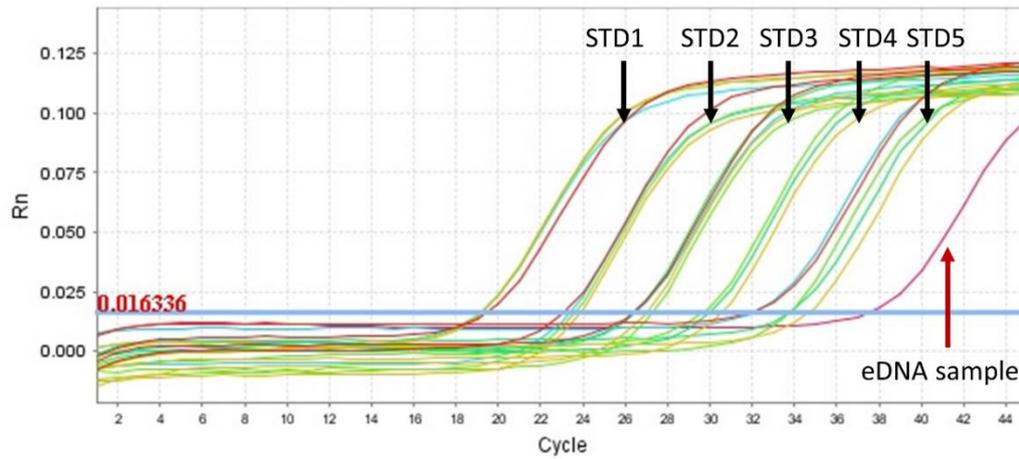
Figure S2. Amplification plot showing positive detection (red arrow) of Arctic charr DNA in one PCR replicate of a Kielder Water eDNA sample. Black arrows indicate qPCR standards (5 points and 5 replicates).