MR. JIANLONG LI (Orcid ID : 0000-0002-0302-0061)

Article type : Resource Article

The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding

Running title: Effect of filters on eDNA metabarcoding

Jianlong Li[1]*, Lori-Jayne Lawson Handley[1], Daniel S. Read[2], Bernd Hänfling[1]

1 Evolutionary and Environmental Genomics Group (@EvoHull), School of Environmental Sciences, University of Hull (UoH), Cottingham Road, Hull, HU6 7RX, UK

2 Centre for Ecology & Hydrology (CEH), Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

*Corresponding author:

haikuilee@gmail.com (J.L.)

Abstract

Environmental DNA (eDNA) is a promising tool for rapid and non-invasive biodiversity monitoring. eDNA density is low in environmental samples, and a capture method, such as filtration, is often required to concentrate eDNA for downstream analyses. In this study, six treatments, with differing filter types and pore sizes for eDNA capture, were compared for their efficiency and accuracy to assess fish community structure with known fish abundance and biomass via eDNA metabarcoding. Our results showed that different filters (with the exception of 20 µm large-pore filters) were broadly consistent in their DNA capture ability. The 0.45 µm filters performed the best in terms of total DNA yield, probability of species detection, repeatability within pond and consistency between ponds. However performance of 0.45 µm filters were only marginally better than for 0.8 µm filters, while filtration time was significantly longer. Given this trade-off, the 0.8 µm filter is the optimal pore size of membrane filter for turbid, eutrophic and high fish density ponds analysed here. The 0.45 µm Sterivex enclosed filters performed reasonably well and are suitable in situations where on-site filtration is required. Finally, pre-filters are applied only if absolutely essential for reducing the filtration time or increasing the throughput volume of the capture filters. In summary, we found encouraging similarity in the results obtained from different filtration methods, but the optimal pore size of filter or filter type might strongly depend on the water type under study.

Keywords: eDNA method development, fish monitoring, pre-filtration, lentic systems

Introduction

The analysis of environmental DNA (eDNA) is a non-invasive genetic method to detect the presence of organisms, including cryptic taxa, that takes advantage of intracellular or extra-organismal DNA in the environment (Lawson Handley 2015; Thomsen & Willerslev 2015; Goldberg *et al.* 2016). Generally, eDNA density is low in environmental samples, and a capture method is therefore required to concentrate eDNA for downstream analyses. The two main approaches to capture eDNA in aquatic environments are precipitation and filtration.

Capturing eDNA through precipitation entails adding ethanol or isopropanol with sodium acetate to water samples (Dejean *et al.* 2011; Foote *et al.* 2012; Doi *et al.* 2017). Samples can be preserved quickly and easily in the field using such an approach, but it is only feasible for small volumes of water (<30 mL), which could reduce the probability of detection, particularly of rare species (Deiner *et al.* 2015; Eichmiller *et al.* 2016). Therefore, most recent studies have used filtration-based methods, which can process larger volumes of typically 250 mL to 5 L, or even up to 45 L (Civade *et al.* 2016). Previous studies have used a wide range of filter types (e.g. different membrane materials and pore sizes) and approaches (e.g. on-site or in laboratory) to filtration. On-site filtration followed by immediate preservation theoretically enhances DNA integrity and is critical for some remote field surveys where access to laboratory facilities is not available. Enclosed filters such Sterivex units (Millipore) or Nalgene analytical test filter funnels (Thermo Fisher Scientific), in combination with a portable peristaltic or hand-driven pump are popular protocols for the capture of eDNA in the field (Keskin 2014; Bergman *et al.* 2016; Wilcox *et al.* 2016; Spens *et al.* 2017). However, a larger number of water samples can be filtered simultaneously in a laboratory setting, which reduces the processing time. Four main types of membrane filter (so-called "open filters") are commonly used in the laboratory set-ups of freshwater studies: (1) 0.45 μm cellulose nitrate

(CN) filters (e.g. Goldberg *et al.* 2011; Pilliod *et al.* 2013), (2) 0.45 µm nylon filters (e.g. Thomsen *et al.* 2012), (3) 0.7 or 1.5 µm glass fibre (GF) filters (e.g. Wilcox *et al.* 2013; Miya *et al.* 2015) and (4) 1.2 µm polycarbonate (PC) filters (e.g. Egan *et al.* 2015).

The suitability of various pore sizes of filter to capture eDNA may be heavily influenced by the heterogeneous nature of aquatic ecosystems. Suspended particulate matter (SPM, e.g. organic matter and sediment) can quickly block 0.2 or 0.45 µm filters (Minamoto *et al.* 2016; Shaw *et al.* 2016), which will severely prolong filtration time and potentially increase concentration of PCR inhibitors (Tsai & Olson 1992; McKee *et al.* 2015). For highly turbid water such as ponds or tropical freshwater ecosystems, even 3 µm PC filters are easily blocked (Minamoto *et al.* 2016; Robson *et al.* 2016). Most previous studies that have investigated the impact of different types and pore sizes of filter on DNA quantity, have focussed on individual target species using real-time quantitative PCR (qPCR) (e.g. Eichmiller *et al.* 2016; Lacoursiere-Roussel *et al.* 2016; Minamoto *et al.* 2016; Robson *et al.* 2016).

Recently, eDNA-based metabarcoding using High-Throughput Sequencing (HTS) has emerged as a powerful tool to monitor entire aquatic communities (e.g. Deiner *et al.* 2016; Hänfling *et al.* 2016; Port *et al.* 2016; Valentini *et al.* 2016). To our knowledge, few previous studies have investigated if and how the choice filtration method impacts on estimates of fish community composition. The preliminary results of Miya *et al.* (2016) showed that the number of detected fish species was significantly higher when using enclosed 0.45 µm polyvinylidene difluoride (PVDF) filters compared to 0.7 µm GF filters, although different filtration systems and extraction methods were used in each case. Djurhuus *et al.* (2017) found that different filter membrane materials (0.2 µm PC, CN, polyethersulfone "PES", and PVDF) and extraction methods did not affect estimates of species richness and community composition across multiple trophic levels. Majaneva *et al.* (2018) indicated that 0.45 µm

MCE filters (described as CN filters in the study) represented the community composition of metazoan more consistently than 0.2 µm PES filters, while the effect of using 12 µm filters as pre-filters remained ambiguous.

The aim of the present study was to further investigate the impact of different filters on eDNA capture and community diversity estimation through eDNA metabarcoding. Specifically, we compared different pore sizes of membrane filter, different types of filter ("open filters" and "enclosed filters"), and the impact of pre-filtration. We evaluated the effect on filtration time, total eDNA recovered, probability of species detection, repeatability, and the relationship between read counts and known fish abundance or biomass in four fish ponds with differing assemblages.

Materials and Methods

Study site and water sampling

This study was carried out at four artificial stock ponds (E1-E4) at the National Coarse Fish Rearing Unit (Nottingham, UK), run by the UK Environment Agency. The size of each pond is 5100 m^2 (60 m × 85 m) and the depth is 1 ~ 1.5 m. Generally, these ponds are used to rear approximately one-year-old common British coarse fish from June to January before they are used in stocking programmes for conservation purposes or recreational fishing. All fish were measured and weighed before stocking in the ponds on 15th June 2015 and after harvesting on 18th January 2016. Fish abundance and biomass at the time of water sampling in August 2015 were estimated, assuming that death and growth curves of these fish are linear (Figs. S1 & S2, Supporting information). The fish stock information in August 2015 is shown in Table 1.

Water sampling was carried out on 6th August 2015. The dissolved oxygen (DO) concentration was similar between ponds (Mean±SD, 7.9±0.8 mg/L). For each pond, 12 water samples were collected at evenly distributed points around the shore. A 1 L sterile bottle was used to collect water at each point just below the surface, and then the water was pooled into a 12.5 L sterile water container. After inverting and shaking the collection container, the water was then subsampled with 25 Gosselin 500 mL sterile plastic bottles. All samples were stored in cool boxes, transferred to the eDNA laboratory at University of Hull (UoH) within 2 hours and refrigerated until filtration.

eDNA capture treatments

Six filtration-based eDNA capture treatments were used for each pond. These treatments were: (1) "0.45MCE": 0.45 µm mixed cellulose acetate and nitrate (also known as mixed cellulose ester or "MCE") filters, 47 mm diameter (Whatman); (2) "0.8MCE": 0.8 µm MCE filters, 47 mm diameter (Whatman); (3) "1.2MCE": 1.2 µm MCE filters, 50 mm diameter (Whatman); (4) "0.45Sterivex": 0.45 µm Sterivex-HV PVDF units (Millipore); (5) "PF_0.45MCE": 0.45 µm MCE filters, 47 mm diameter (Whatman) after pre-filtration with 20 µm qualitative cellulose filters, Grade 4 (Whatman); and (6) "PF": the pre-filters used in treatment 5. Each treatment was replicated five times, filtering 300 mL water each time, resulting in a total of 120 replicates. These treatments were used to measure three different effects: pore sizes (0.45MCE, 0.8MCE and 1.2MCE), filter types (0.45MCE and 0.45Sterivex) and pre-filtration (0.45MCE and PF_0.45MCE) (Fig. 1).

To reduce cross-contamination, the samples from individual ponds were filtered separately in order of pond E1 to E4. For each replicate (apart from the "0.45Sterivex" treatment), 300 mL water was filtered using Nalgene filtration units (Thermo Fisher Scientific) in combination with a vacuum pump (15~20 in. Hg, Pall Corporation). For each pond, the same

filtration unit was used for the all five replicates of the same capture treatment. The filtration units were cleaned with 10% v/v commercial bleach solution and 5% v/v microsol detergent (Anachem, UK), and then rinsed thoroughly with deionised water after each filtration to prevent cross-contamination. Filtration blanks (n=5) with 300 mL deionised water were run before the first filtration and after every wash run in order to test for possible contamination at the filtration stage. For the "0.45Sterivex" treatment, 300 mL water was directly filtered with 0.45 µm Sterivex units in combination with a vacuum pump (15~20 in. Hg, Pall Corporation). All samples were filtered within 24 hours of collection in a dedicated eDNA filtration laboratory at UoH.

After filtration, all membrane filters were placed into 50 mm sterile petri dishes sealed with parafilm, while Sterivex units were closed with inlet and outlet caps. All samples were stored in a freezer at -20^{o}C until DNA extraction. DNA extraction was carried out using the PowerWater (Sterivex) DNA Isolation Kits (MoBio Laboratories Inc., now Qiagen) following the manufacturer's protocol. Total DNA concentration was quantified using a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific) after extraction.

Library preparation and sequencing

Extracted DNA samples were PCR-amplified targeting a 106 bp vertebrate-specific fragment of the mitochondrial 12S rRNA region (Riaz *et al.* 2011) following a one-step library preparation protocol (Kozich *et al.* 2013) with amplification primers that include PCR primers, indices and flow cell adapters. Previous studies showed that this fragment has a low false negative rate in both marine mesocosm and coastal ecosystem eDNA metabarcoding studies of bony fishes (Kelly *et al.* 2014; Port *et al.* 2016). We also previously tested this fragment *in vitro* on 22 common freshwater fish species and *in situ* on three deep lakes in the English Lake District, and demonstrated their suitability for eDNA metabarcoding of UK lake fish communities (Hänfling *et al.* 2016).

All PCRs were set up in a PCR workstation in our dedicated eDNA laboratory to minimize the risk of contamination. All samples (n=120) together with five filtration and extraction controls, five no-template PCR controls and five positive PCR controls (the Eastern Happy, *Astatotilapia calliptera,* a cichlid from Lake Malawi, which is not present in the UK) were included in the Illumina MiSeq library construction and sequencing (n=135). PCR reactions were carried out in 25 μL volumes with the MyTaq HS Red Mix PCR Kit (Bioline) containing: 1X Master Mix, 0.5 μM of each tagged primer and 2.5 μL template DNA. Eight-strip PCR tubes with individually attached lids and mineral oil (Sigma-Aldrich) were used to reduce cross-contamination between samples. PCRs were performed on an Applied Biosystems Veriti thermal cycler with the following profile: 98 °C for 5 min, 35 cycles of 98 °C for 10 sec, 58 °C for 20 sec and 72 °C for 30 sec, followed by a final elongation step at 72 °C for 7 min. Three PCR technical replicates were performed for each sample then pooled to minimize bias in individual PCRs.

PCR products were purified and normalized using the SequalPrep Normalization Plate Kit (Invitrogen) and subsequently pooled in equal volume (i.e. 5 μL per sample). The pooled library was further purified using the QIAquick Gel Extraction Kit (Qiagen) and resuspended in 20 μL elution buffer. The library concentration was then quantified by Qubit v3.0 using the dsDNA HS Assay Kit (Thermo Fisher Scientific). The pooled library was adjusted to 2 nM and denatured following the Illumina MiSeq library denaturation and dilution guide. Because of the low fish diversity in the ponds, the final 10 pM denatured library was mixed with 30% PhiX control to improve the diversity of the library. The library was sequenced on an Illumina MiSeq platform using the MiSeq reagent kit v2 (2×250 cycles) at the UoH. The custom sequencing and index primers were added to the appropriate wells of the MiSeq reagent cartridge as described by Kozich *et al.* (2013).

Data analysis

Bioinformatics analysis

Raw read data from Illumina MiSeq sequencing have been submitted to NCBI (BioProject: PRJNA414952; BioSample accession: SAMN07811461~SAMN07811580; Sequence Read Archive accessions: SRR6189420~SRR6189539). Bioinformatics analysis was implemented following a custom reproducible metabarcoding pipeline (metaBEAT v0.97.8) with a custom-made 12S rRNA reference database as described in our previous study (Hänfling *et al.* 2016). The maximum likelihood phylogenetic tree of the all 12S rRNA sequences from the custom reference database is shown in Fig. S3 (Supporting information). Sequences for which the best BLAST hit had a bit score below 80 or had less than 100% identity to any sequence in the curated database were considered non-target sequences. To assure full reproducibility of our bioinformatics analysis, the up to date (May 2017) custom reference database and the Jupyter notebook for data processing have been deposited in an additional dedicated GitHub repository (https://github.com/HullUni-bioinformatics/Li_et_al_2018_eDNA_filtration).

Criteria for reducing false positives and quality control

Filtered data were summarized into the number of sequence reads per species (hereon referred to as read counts) for downstream analyses (Appendix S1, Supporting information). We applied two criteria to reduce the possibility of false positives. (1) The low-frequency noise threshold (proportion of positive species read counts of all read counts in the real sample) was set to filter some high-quality annotated reads passing the previous filtering steps that have high-confidence BLAST matches but may be inaccurate due to potential low-level contamination during the library construction process (De Barba *et al.* 2014; Hänfling *et al.* 2016; Port *et al.* 2016). The low-frequency noise threshold was set to 0.001 in this study as determined empirically in Hänfling *et al.* (2016); therefore, all taxonomic assignments

with frequency below this threshold were omitted from further downstream analysis. (2) After the low-frequency noise threshold was applied, remaining taxonomic assignments of taxa that were not stocked in the ponds (i.e. *Salmo trutta*, *Alburnus alburnus* and *Gobio gobio*) were also treated as false positives and excluded.

Samples were excluded from the analysis because they performed poorly in terms of PCR and sequencing depth due to low DNA concentrations. Two samples (T3-1-3 and T2-2-3) showed extremely low levels of DNA concentration and failed PCR. One sample (T4-1-3) had only slightly reduced DNA concentration but consistently produced poor results during PCR which resulted in no read count assigned to fish (Fig. 2; Fig. S4, Supporting information).

Similarity and statistical analyses

All similarity and statistical analyses were performed in R v3.3.2 (R_Core_Team 2016) and graphs were plotted using *ggplot2* v2.2.1 (Wickham & Chang 2016).

To better quantify the heterogeneity between filtration replicates, the Horn similarity index was calculated based on species relative abundance using *SpadeR* v0.1.1 (Chao *et al.* 2016) with the function *SimilarityMult*. To investigate effects of different capture treatments on fish communities, non-metric multidimensional scaling (NMDS) allied with analysis of similarities (ANOSIM) were performed using the abundance-based Bray-Curtis dissimilarity index with the function *metaMDS* and *anosim* respectively in *Vegan* v2.4-4 (Oksanen *et al.* 2017). The treatment with high repeatability should have high mean Horn index and low variation in NMDS ordination. The ANOSIM statistic R is based on the difference of mean ranks between treatments and within treatments.

Two-way analysis of variance (ANOVA) was conducted to test the interaction between four ponds and six treatments for filtration time, total DNA yield, probability of species detection, Horn index and correlation coefficient between read counts and abundance or

biomass after square-root or Tukey's ladder of powers transformation. Kruskal-Wallis one-way ANOVA with Dunn's test was conducted to test differences between the capture treatments for filtration time and Horn index. ANOVA with Tukey's test was conducted to test differences between the capture treatments for total DNA yield. The significance of linear correlations between read counts and abundance or biomass was evaluated by calculating the Pearson's product-moment correlation coefficient.

The full R script is available on the GitHub repository (https://github.com/HullUni-bioinformatics/Li_et_al_2018_eDNA_filtration/tree/master/R_script).

Results

Filtration time

The filtration time across all treatments and ponds varied from 3 to 120 min (Fig. 3). There were significant effects of "treatment", "pond", the "interaction" between ponds and treatments across the entire data set (Table 2, Global), and when comparing different treatments under specific aims (Table 2). The average filtration time differed considerably among the four ponds under the same filtration treatment, suggesting that SPM content varied among ponds (Table S1, Supporting information). In relation to the specific comparisons: the filtration time decreased on average by 19.88±14.17 min (Mean±SD) when the pore size increased from 0.45 to 0.8 µm and by 5.68±5.98 min (Mean±SD) when the pore size increased from 0.8 to 1.2 µm. Overall, filtration time significantly decreased with increasing pore size, but the pattern was complex since significant interactions between treatments and ponds were observed (Table 2, Pore sizes). Individual *post hoc* tests showed that not all pairwise comparisons among pore sizes were significant (e.g. pond E4). Filtration time was on average 18.00±6.48 min (Mean±SD) longer using the "0.45Sterivex" compared to the

"0.45MCE". This pattern was also seen in three out of the four ponds when looked at individually but none of the *post hoc* tests within ponds were significant (Fig. 3). Across the four ponds, it was possible to filter 300 mL water in around 4 min using pre-filters themselves (Fig. 3; Table S1, Supporting information). Filtration time decreased on average by 27.00±13.87 min (Mean±SD) when comparing the 0.45 μm filters after pre-filtration ("PF_0.45MCE") to those without pre-filtration ("0.45MCE"); and this significant trend was observed in ponds E1 and E3 (Fig. 3A, C).

DNA yield

The DNA concentration across all treatments and ponds ranged from 1.15 to 119.70 ng/μL (Fig. 2). There were significant effects of "treatment", "pond", the "interaction" between ponds and treatments across the entire data set (Table 2, Global). In relation to the specific comparisons: there was no significant effect of different pore sizes of filter (Table 2, Pore sizes, P=0.07). Comparing the "0.45Sterivex" and the "0.45MCE", there were significant effects of "treatment" and "pond" (Table 2, Filter types). Individual *post hoc* tests showed that there was no significant difference between using the "0.45Sterivex" and the "0.45MCE" treatments from ponds E1 to E3, but the total DNA yield recovered from the "0.45Sterivex" was significantly lower than the "0.45MCE" in pond E4 (Fig. 2D). The average DNA yield recovered from the pre-filters themselves ("PF") was the lowest of the six filtration treatments (Table S1, Mean±SD, 16.65±9.85 ng/μL, Supporting information). After pre-filtration, the "PF_0.45MCE" still recovered 73.27±10.56% (Mean±SD) total eDNA; hence only 26.73±10.56% (Mean±SD) of the total eDNA remained on the 20 μm pre-filters. There were significant effects of "treatment" and "pond" between the "0.45MCE" and the "PF_0.45MCE" (Table 2, Pre-filtration). Individual *post hoc* tests showed that the total DNA

yield recovered from the "0.45MCE" was significantly higher than the "PF_0.45MCE" in pond E4 only (Fig. 2D).

Probability of species detection

All eight stocked species (*Abramis brama*, *Barbus barbus*, *Carassius carassius*, *Squalius cephalus*, *Leuciscus leuciscus*, *Rutilus rutilus*, *Scardinius erythrophthalmus* and *Tinca tinca*) were detected in this study (Fig. 4). The rarest species in ponds E1 and E2 was *A. brama.* This species was not detected in pond E2 with any treatment, but it was detected with "0.45Sterivex" in pond E1. *Rutilus rutilus* was not detected using the pre-filters ("PF") in pond E2 (Fig. 4). In ponds E3 and E4, all stocked species were detected by all of the treatments (Fig. 4C, D). There were significant effects of "treatment" and "pond" across the entire data set, but there was no significant difference of "interaction" between ponds and treatments (Table 2, Global). In relation to the specific comparisons: there was no significant difference when comparing different filter pore sizes (Table 2, Pore sizes, P=0.16), and filtration with and without pre-filters (Table 2, Pre-filtration, P=0.43). The Sterivex units ("0.45Sterivex") performed slightly better than the "0.45MCE" in terms of probability of species detection (Table 2, Filter types, $P<0.05$). The average probability of species detection was the lowest using the pre-filters themselves ("PF") of the six filtration treatments (0.64±0.27, Table S1, Supporting information).

Variation between filtration replicates

Overall, there was considerable variation in species composition among individual filtration replicates within ponds (Fig. 5A1, B1, C1, D1; Fig. S5, Supporting information). In terms of Horn index (similarity between replicates), there were significant effects of

"treatment", "pond", the "interaction" between ponds and treatments across the entire data set (Table 2, Global), and when comparing different treatments under specific aims (Table 2). The NMDS showed a high degree of overlap between the six capture treatments across four ponds (Fig. 5A2, B2, C2, D2) indicating that different filtration treatments yielded broadly similar community composition estimates. Notable exceptions to this pattern were the pre-filters ("PF") and in some ponds (e.g. ponds E1 & E2) "PF_0.45MCE", where individual replicates were more widely scattered and often outside the ellipses of other treatments. In the ANOSIM test, the average values of the R statistic in global tests with all treatments were low (Table S2, Mean±SD, 0.15±0.03, Supporting information), which showed that there was no obvious difference between treatments; and the P values suggesting that the variation was attributed to filtration replicates instead of treatments (Table S2, Mean±SD, P=0.03±0.02, Supporting information).

In relation to the specific comparisons: overall, Horn index significantly decreased with increasing pore size, but the pattern was complex since significant interactions between treatments and ponds were observed (Table 2, Pore sizes). Individual *post hoc* tests showed that not all pairwise comparisons among pore sizes were significant (e.g. pond E2). The NMDS analysis showed that there was only clear discrimination between the "0.45MCE" and the "0.8MCE" in pond E1 (Fig. 5A2; Table S2, ANOSIM: R=0.52, P=0.01, Supporting information). There was greater variation among the "0.45Sterivex" replicates compared to the "0.45MCE" replicates (Fig. 5). The community similarity of the "0.45Sterivex" was significantly lower than the "0.45MCE" across four ponds (Table 2, Filter types; Fig. 5A1, B1, C1, D1). The NMDS ordination showed that significant difference was observed between the "0.45Sterivex" and the "0.45MCE" in ponds E3 (Fig. 5C2; Table S2, ANOSIM: R=0.64, P=0.02, Supporting information) and E4 (Fig. 5D2; Table S2, ANOSIM: R=0.30, P=0.02, Supporting information). Greater variance between replicates was observed for the pre-filters

("PF") themselves compared to other treatments (Fig. 5). Repeatability was similar for the 0.45 μm filters when using pre-filters ("PF_0.45MCE") and without using pre-filters ("0.45MCE"), except in pond E1 where the Horn index was significantly lower for "PF_0.45MCE" than "0.45MCE" (Fig. 5A1). The NMDS ordination showed that there was no significant difference between the "PF_0.45MCE" and the "0.45MCE" across four ponds (Fig. 5A2, B2, C2, D2; Table S2, Mean±SD, ANOSIM: R=0.07±0.06, P=0.26±0.12, Supporting information).

Correlations between read counts and fish abundance or biomass

There were consistent, positive correlations between average read counts of five replicates and fish abundance or biomass across the six treatments and four ponds (Fig. 6; Fig. S6, Supporting information). There was no significant effect of "treatment", or "interaction" between ponds and treatments, on correlations between read counts and abundance or biomass across the entire data set (Table 2, Global). In relation to the specific comparisons: overall, there were significant effects of different pore sizes of filter (Table 2, Pore sizes). Individual *post hoc* tests showed that a significant difference in correlations between read counts and abundance or biomass was only observed between "0.45MCE" and "1.2MCE" treatments, and the 1.2 μm MCE filters performed better than 0.45 μm MCE filters. There was no significant effect on correlations between read counts and abundance or biomass between "0.45Sterivex" and "0.45MCE" treatments (Table 2, Filter types), and filtration with and without pre-filtration (Table 2, Pre-filtration).

Discussion

Optimal pore size of membrane filter

Turner *et al.* (2014) previously determined that aqueous eDNA particles from common carp (*Cyprinus carpio*) ranged between < 0.2 and > 180 µm and therefore recommended 0.2 µm pore size filters for optimal capture of common carp eDNA. In a pilot study, we observed that this pore size of filter quickly led to clogging; therefore we compared three pore sizes (0.45, 0.8 and 1.2 µm) of membrane filter.

Our study demonstrated that the filter pore size had considerable impact on filtration time. When changing from 0.45 to 0.8 µm filters, on average, 36% filtration time was saved, whereas only 15% filtration time was saved increasing pore size from 0.8 to 1.2 µm. This result supports previous studies (Turner *et al.* 2014; Eichmiller *et al.* 2016; Minamoto *et al.* 2016) indicating that the smaller pore size of filters were more likely to clog and increase filtration time. However, different pore sizes did not affect the amount of total eDNA recovered and probability of species detection. The similarity among filtration replicates decreased with increasing pore size; and the repeatability among filtration replicates using the 0.45 µm MCE filters was the highest compared to the other pore sizes of filter. This in turns indicates that stochastic sampling effects can be minimised by using smaller pore size of filters. After pooling that data from all five replicates consistently positive relationships were found between read counts and fish abundance or biomass, although correlations were not always statistically significant. The 0.8 µm and 1.2 µm MCE filters performed better than 0.45 µm MCE filters in terms of correlations between read counts and fish abundance or biomass. In contrast, Eichmiller *et al.* (2016) found that different pore sizes (0.2, 0.6, 1.0 and 5.0 µm) of PC filter affected the slope of the *C. carpio* biomass/eDNA copies relationship; and 0.2~0.6 µm filters were optimal for biomass quantification in the laboratory. Turner *et al.*

(2014) showed that PC filters have relatively uniform sized pores, in contrast, the MCE filters are less uniform and more likely to retain particles by entrapment. The structural difference between PC filters and MCE filters could explain why our results are different from Eichmiller *et al.* (2016). Previous studies have also demonstrated that filter materials can also drastically affect the recovery of eDNA (Liang & Keeley 2013; Renshaw *et al.* 2015; Hinlo *et al.* 2017). The other potential reason for difference between studies could be that previous studies were based on target species detection via qPCR assays, comparing absolute DNA concentrations across samples, as opposed to metabarcoding of the whole community comparing relative sequencing read counts in the current study. In support of this, Djurhuus *et al.* (2017) found that different filter materials did not result in different richness and community composition based on metabarcoding.

The 0.45 µm MCE filters performed the best among the six filtration treatments in terms of DNA yield, repeatability within pond and consistency between ponds. However, filtration time was significantly longer for the 0.45 µm MCE filters than the 0.8 µm MCE filters. The correlations between read counts and fish abundance or biomass recovered by the 0.8 µm MCE filters were slightly better than those of the 0.45 µm MCE filters even though there was no significant difference between the treatments. Therefore, the 0.8 µm MCE filters appear to provide a reasonable balance between filtration time and quantification efficacy in this study and may be optimal in turbid, eutrophic, high fish density water bodies, whereas 0.45 µm MCE filters may be more suitable to clearer waters (Fig. 1).

Performance of enclosed (Sterivex) filters

Previous studies showed that filtration using enclosed Sterivex units is an effective protocol for capturing target species DNA with qPCR assays (Keskin 2014; Bergman *et al.* 2016; Spens *et al.* 2017). To our knowledge, Spens *et al.* (2017) is the only published study

comparing Sterivex units with membrane filters using qPCR. Here, we directly compared the performance of MCE filters and Sterivex units of the same pore size via metabarcoding.

On average, filtration time using the Sterivex units increased 18 min per sample compared to using 0.45 µm MCE filters. This difference is not due to vacuum pumps as the same pump was used for both filter types. However, Spens *et al.* (2017) observed that 1 L clear lake water can be filtered through 0.22 µm Sterivex units in around 10 min using 50 mL syringes comparing to 0.45 µm MCE filters (described as CN filters in the study) in 15~30 min using a vacuum pump. To minimize filtration time, we therefore recommend that Sterivex units are used together with prepacked sterile syringes in situations where on-site filtration is required (Fig. 1). With respect to DNA yield, the 0.45 µm Sterivex filters recovered slightly less DNA than the 0.45 µm MCE filters. The Horn index and NMDS ordination showed there was a greater variation among the 0.45 µm Sterivex replicates compared to the 0.45 µm MCE replicates. However, the correlations between read counts and fish biomass or abundance were not significantly different between the treatments when all data were pooled. Therefore, 0.45 µm Sterivex units can be considered an efficient eDNA capture method for metabarcoding.

Efficiency and impact of pre-filtration

The water from Calverton fish ponds is turbid and eutrophic, with high levels of algae. Our pilot study showed that a small amount of water (i.e. 250 mL) could be filtered through 1.2 µm filters before clogging. This is considerably less than previous metabarcoding studies in less eutrophic lakes, in which at least 1 L water was filtered (Hänfling *et al.* 2016; Port *et al.* 2016) and reduced sample volumes could potentially impact rare species detection. Pre-filtration could potentially help to prevent clogging, substantially reduce filtration time, and reduce the capture of unwanted SPM and PCR inhibitors. We therefore investigated the

impact of pre-filtration by comparing results from 0.45 µm MCE filters with and without passing through 20 µm pre-filters, as well as the analysing pre-filters themselves.

Across the four ponds, it was possible to filter 300 mL water in around 4 min using the pre-filters themselves. The pre-filtering step reduced the filtration time through the 0.45 µm MCE filters by approximately 50%, resulting in a considerable overall time saving per sample. This could be an important consideration when eutrophic habitat or water with high sediment content is sampled. After pre-filtration, 73.27% total eDNA was recovered on the 0.45 µm MCE filters (with a corresponding 26.73% total eDNA remained on pre-filters). Pre-filtration followed by capture onto 0.45 µm MCE filters did not result in significantly different probability of species detection, repeatability between filtration replicates, and correlations between read counts and fish biomass or abundance when compared to other treatments. However, Majaneva *et al.* (2018) demonstrated that pre-filtration (12 µm pre-filters with 0.45 µm filters), could potentially reduce the number of detected metazoan taxa, although it recovered higher diversity index values and more consistent community composition.

In terms of the pre-filters themselves, the overall probability of species detection (0.64±0.27) was lower than other membrane filters, and greater variance between replicates was observed compared to other treatments. Similar results were found by Robson *et al.* (2016), who showed that 2 L water samples can be filtered in less than 3 min using 20 µm filters, but a 0.57 probability of single species detection was achieved compared to 1.00 probability using 3 µm PC filters.

Our results indicate that pre-filtration with 20 µm filters could prevent SPM from clogging finer filters without affecting metabarcoding results but that the pre-filters themselves are not suitable for metabarcoding due to the potential of reduced total DNA yield, probability of species detection and repeatability. Despite the advantages of pre-filtration demonstrated here,

it should be noted that there is a drawback of pre-filtration in terms of more handling, which could increase the opportunity for contamination (Turner *et al.* 2014). Thus, we recommend pre-filters are applied only if absolutely essential for reducing the filtration time or increasing the throughput volume of the capture filters (Fig. 1).

Conclusion

This study demonstrate that the DNA yield, probability of species detection, and correlations between abundance/biomass and read counts are encouragingly comparable between different filter types (0.45 MCE filters and 0.45 Sterivex units) and pore sizes (0.45, 0.8 and 1.2 μm). Therefore, eDNA metabarcoding results seem quite robust to the choice of the filtration method when a sufficient number of replicates is carried out. We note, however, that the suitability of various pore sizes of filter to capture eDNA is likely to be heavily influenced by the heterogeneous nature of water bodies. For turbid, eutrophic, high fish density ponds, such as those studied here, 0.8 μm MCE filters provide the optimal trade-off between rapid filtration time and probability of species detection, but smaller pore sizes of filter may be more suitable for clearer, low species density conditions. Further study of the impact of heterogeneity (in terms of SPM, biochemical oxygen demand "BOD", chemical oxygen demand "COD", dissolved oxygen "DO", pH, water colour etc.) between water bodies on eDNA capture is required. Finally, we report high variation among filtration replicates, which is consistent with Lanzén *et al.* (2017) who indicated that technical replicates of DNA extraction can improve diversity and compositional dissimilarity. Spatial heterogeneity of eDNA within water bodies has also been reported in several studies (e.g. Jerde *et al.* 2011; Pilliod *et al.* 2013; Civade *et al.* 2016; Hänfling *et al.* 2016). Future studies, for example incorporating species occupancy models for imperfect species detection (Pilliod

et al. 2013; Schmidt *et al.* 2013; Hänfling *et al.* 2016; Valentini *et al.* 2016), are needed to further investigate the multiple opportunities for heterogeneity encountered in eDNA studies.

Acknowledgments

This work is part of PhD project of J.L., who is supported by University of Hull and China Scholarship Council. We are particularly grateful to Mr Alan Henshaw and other staff of Calverton fish farm of UK Environment Agency for sampling and providing fish stock information, and Drs Christoph Hahn and Amir Szitenberg for helping on the bioinformatics analysis. Furthermore, Dr James Kitson provided excellent advice on library preparation. We would like to thank three anonymous reviewers for constructive criticism on the initial submission which helped to strengthen the manuscript, and Lynsey Harper and Cristina Di Muri for helping proofreading and discussion on DNA quantification.

References

Bergman PS, Schumer G, Blankenship S, Campbell E (2016) Detection of adult green sturgeon using environmental DNA analysis. *PLoS ONE*, **11**, e0153500.

Chao A, Ma KH, Hsieh TC, Chiu CH (2016) SpadeR: species-richness prediction and diversity estimation with R): an R package in CRAN. Available from https://CRAN.R-project.org/package=SpadeR.

Civade R, Dejean T, Valentini A *et al.* (2016) Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PloS ONE*, **11**, e0157366.

De Barba M, Miquel C, Boyer F *et al.* (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources*, **14**, 306-323.

Deiner K, Fronhofer EA, Machler E, Walser JC, Altermatt F (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, **7**, 1-9.

Deiner K, Walser JC, Machler E, Altermatt F (2015) Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*, **183**, 53-63.

Dejean T, Valentini A, Duparc A *et al.* (2011) Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*, **6**, e23398.

Djurhuus A, Port J, Closek CJ *et al.* (2017) Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Frontiers in Marine Science*, **4**, 314.

Doi H, Uchii K, Matsuhashi S *et al.* (2017) Isopropanol precipitation method for collecting fish environmental DNA. *Limnology and Oceanography-Methods*, **15**, 212-218.

Egan SP, Grey E, Olds B *et al.* (2015) Rapid molecular detection of invasive species in ballast and harbor water by integrating environmental DNA and light transmission spectroscopy. *Environmental Science & Technology*, **49**, 4113-4121.

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

Foote AD, Thomsen PF, Sveegaard S *et al.* (2012) Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS ONE*, **7**, e41781.

Goldberg CS, Pilliod DS, Arkle RS, Waits LP (2011) Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. *PLoS ONE*, **6**, e22746.

Goldberg CS, Turner CR, Deiner K *et al.* (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, **7**, 1299-1307.

Hänfling B, Lawson Handley L, Read DS *et al.* (2016) Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, **25**, 3101-3119.

Hinlo R, Gleeson D, Lintermans M, Furlan E (2017) Methods to maximise recovery of environmental DNA from water samples. *PLoS ONE*, **12**, e0179251.

Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, **4**, 150-157.

Kelly RP, Port JA, Yamahara KM, Crowder LB (2014) Using environmental DNA to census marine fishes in a large mesocosm. *PLoS ONE*, **9**, e86175.

Keskin E (2014) Detection of invasive freshwater fish species using environmental DNA survey. *Biochemical Systematics and Ecology*, **56**, 68-74.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, **79**, 5112-5120.

Lacoursiere-Roussel A, Rosabal M, Bernatchez L (2016) Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. *Molecular Ecology Resources*, **16**, 1401-1414.

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

Lawson Handley L (2015) How will the 'molecular revolution' contribute to biological recording? *Biological Journal of the Linnean Society*, **115**, 750-766.

Liang Z, Keeley A (2013) Filtration recovery of extracellular DNA from environmental water samples. *Environmental Science & Technology*, **47**, 9324-9331.

Majaneva M, Diserud OH, Eagle SHC *et al.* (2018) Environmental DNA filtration techniques affect recovered biodiversity. *Scientific Reports*, **8**, 4682.

McKee AM, Spear SF, Pierson TW (2015) The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biological Conservation*, **183**, 70-76.

Minamoto T, Naka T, Moji K, Maruyama A (2016) Techniques for the practical collection of environmental DNA: filter selection, preservation, and extraction. *Limnology*, **17**, 23-32.

Miya M, Minamoto T, Yamanaka H *et al.* (2016) Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. *Journal of Visualized Experiments*, **117**, e54741.

Miya M, Sato Y, Fukunaga T *et al.* (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, **2**, 150088.

Oksanen J, Blanchet FG, Friendly M *et al.* (2017) Vegan: community ecology package. Available from https://CRAN.R-project.org/package=vegan.

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

Port JA, O'Donnell JL, Romero-Maraccini OC *et al.* (2016) Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, **25**, 527-541.

R_Core_Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

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

Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, gkr732.

Robson HL, Noble TH, Saunders RJ *et al.* (2016) Fine-tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems. *Molecular Ecology Resources*, **16**, 922-932.

Schmidt BR, Kery M, Ursenbacher S, Hyman OJ, Collins JP (2013) Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution*, **4**, 646-653.

Shaw JL, Weyrich L, Cooper A (2016) Using environmental (e) DNA sequencing for aquatic biodiversity surveys: a beginner's guide. *Marine and Freshwater Research*.

Spens J, Evans AR, Halfmaerten D *et al.* (2017) Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, **8**, 635-645.

Thomsen PF, Kielgast J, Iversen LL *et al.* (2012) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, **7**, e41732.

Thomsen PF, Willerslev E (2015) Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4-18.

Tsai Y-L, Olson BH (1992) Detection of low numbers of bacterial cells in soils and sediments by polymerase chain reaction. *Applied and Environmental Microbiology*, **58**, 754-757.

Turner CR, Barnes MA, Xu CC *et al.* (2014) Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, **5**, 676-684.

Valentini A, Taberlet P, Miaud C *et al.* (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, **25**, 929-942.

Wickham H, Chang W (2016) ggplot2: create elegant data visualisations using the grammar of graphics. Available from https://CRAN.R-project.org/package=ggplot2

Wilcox TM, McKelvey KS, Young MK *et al.* (2013) Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLoS ONE*, **8**, e59520.

Wilcox TM, McKelvey KS, Young MK *et al.* (2016) Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation*, **194**, 209-216.

Data Accessibility

Raw read data from Illumina MiSeq sequencing have been submitted to NCBI (BioProject: PRJNA414952; BioSample accession: SAMN07811461~SAMN07811580; Sequence Read Archive accessions: SRR6189420~SRR6189539). To assure full reproducibility of our analyses, we have deposited the entire bioinformatics workflow in a dedicated Github repository, which also contains the curated reference database and further supplementary data, such as R scripts (https://github.com/HullUni-bioinformatics/Li_et_al_2018_eDNA_filtration); the repository is permanently archived with Zenodo (https://doi.org/10.5281/zenodo.1228089). Our custom data processing pipeline is available on Github (https://github.com/HullUni-bioinformatics/metaBEAT).

Author Contributions

B.H. and J.L. conceived and designed the study; J.L. performed experiment and bioinformatics analyses; D.S.R. assisted in the design of the workflows for MiSeq library preparation. J.L., L.L.H. and B.H. performed the statistical analyses; J.L. wrote the manuscript; all authors commented the final manuscript.

Figure Captions

Figure 1: Flow chart illustrating selection of eDNA capture, preservation and extraction based on the filtration equipment and aquatic ecosystems of study. "MCE": mixed cellulose acetate and nitrate. Note: Pre-filters are applied only if it substantially reducing the filtration time or increasing the throughput volume of the capture filters. '†' refers to this method was recommended by Spens et al. (2017).

Figure 2: DNA yield recovered from six eDNA capture treatments from four ponds (A-D correspond to ponds E1-E4 respectively). Five replicates under each treatment. Treatments that differ significantly ($P<0.05$) are indicated by the different letters above the bars. "0.45MCE": 0.45 µm mixed cellulose acetate and nitrate (MCE) filters; "0.8MCE": 0.8 µm MCE filters; "1.2MCE": 1.2 µm MCE filters; "0.45Sterivex": 0.45 µm Sterivex-HV enclosed units; "PF_0.45MCE": 0.45 µm MCE filters after 20 µm qualitative cellulose pre-filters, and "PF": 20 µm qualitative cellulose pre-filters. Note: 'Diamonds ◊' show average values and the white dots represent outliers, identified in 'Data analysis' section, are excluded downstream analysis.

Figure 3: Filtration time of six eDNA capture treatments from four ponds (A-D correspond to ponds E1-E4 respectively). Five replicates under each treatment. Treatments that differ significantly ($P<0.05$) are indicated by the different letters in boxplots. Abbreviations of treatments are the same as in Fig. 2. Note: 'Diamonds ◊' show average values and the white dots represent outliers, identified in 'Data analysis' section, are excluded downstream analysis.

Figure 4: Species composition of averaged read counts (number of replicates = 5) using six eDNA capture treatments of eDNA from four ponds (A-D correspond to ponds E1-E4 respectively). Species three letter codes are given in Table 1 and abbreviations of treatments are the same as in Fig. 2. 'Bio' and 'Abu' refer to species composition of fish biomass or abundance calculated based on Table 1, respectively. Note: Replicates identified as outliers are excluded.

Figure 5: Pairwise Horn similarity index (A1-D1) and non-metric multidimensional scaling (NMDS) (A2-D2) based on six eDNA capture treatments from four ponds (A-D correspond to ponds E1-E4 respectively). 'Among' refers to all filtration replicates among treatments within pond (A1-D1). Treatments that differ significantly ($P<0.05$) are indicated by the different letters in boxplots (A1-D1). The ellipses indicate the 50% standard error of each capture method in order to visualise the individual data points (which are not visible at 95%) (A2-D2). Species three letter codes are given in Table 1 and abbreviations of treatments are the same as in Fig. 2. Note: Five replicates under each treatment and replicates identified as outliers are excluded.

Figure 6: Correlations between averaged read counts (number of replicates = 5) and fish abundance using six eDNA capture treatments from four ponds (A-D correspond to ponds E1-E4 respectively). Abbreviations of treatments are the same as in Fig. 2. Note: Replicates identified as outliers are excluded.

Table 1: Fish stock information of four experiment ponds at the National Coarse Fish Rearing Unit.

Pond	Species			August 2015	
	Scientific name	Common name	Code	Abundance	Biomass(kg)
E1	*Rutilus rutilus*	Roach	ROA	33515	199.7
E1	*Barbus barbus*	Barbel	BAR	9695	118.8
E1	*Squalius cephalus*	Chub	CHU	14943	445.2
E1	*Abramis brama*	Common bream	BRE	500	7.1
E1	*Tinca tinca*	Tench	TEN	944	10.9
E1	*Carassius carassius*	Crucian Carp	CAR	489	10.2
E2	*Rutilus rutilus*	Roach	ROA	4730	52.4
E2	*Leuciscus leuciscus*	Dace	DAC	34729	287.0
E2	*Barbus barbus*	Barbel	BAR	9691	295.6
E2	*Abramis brama*	Common bream	BRE	487	4.7
E2	*Carassius carassius*	Crucian Carp	CAR	4910	86.8
E3	*Squalius cephalus*	Chub	CHU	18967	542.6
E3	*Rutilus rutilus*	Roach	ROA	30156	321.2
E3	*Carassius carassius*	Crucian Carp	CAR	3474	58.6
E3	*Tinca tinca*	Tench	TEN	4773	58.2
E4	*Leuciscus leuciscus*	Dace	DAC	29322	248.0
E4	*Barbus barbus*	Barbel	BAR	9508	268.7
E4	*Scardinius erythrophthalmus*	Rudd	RUD	8334	71.1
E4	*Abramis brama*	Common bream	BRE	4962	52.6
E4	*Carassius carassius*	Crucian Carp	CAR	199	17.6
E4	*Tinca tinca*	Tench	TEN	4763	43.5

Note: Full scientific, common names and three letter codes used in figures are given.

Table 2: Two-way analysis of variance (ANOVA) results for filtration time, total DNA yield, species detection probability, correlation with abundance, and correlation with biomass using six eDNA capture treatments across four ponds (E1-E4).

Evaluation criterion	Group	Treatment	Pond	Interaction
Filtration time (min)	Global	F(5,93)=234.96***	F(3,93)=288.44***	F(15,93)=14.35***
	Pore sizes	F(2,46)=47.88***	F(3,46)=173.90***	F(6,46)=4.31**
	Filter types	F(1,31)=12.43**	F(3,31)=61.92***	F(3,31)=5.11**
	Pre-filtration	F(1,32)=123.11***	F(3,32)=169.41***	F(3,32)=4.12*
Total DNA yield (ng/μL)	Global	F(5,93)=42.07***	F(3,93)=24.06***	F(15,93)=2.96***
	Pore sizes	F(2,46)=2.82; P=0.07	F(3,46)=17.61***	F(6,46)=3.46**
	Filter types	F(1,31)=34.00***	F(3,31)=8.63***	F(3,31)=1.09; P=0.36
	Pre-filtration	F(1,32)=8.57**	F(3,32)=4.49**	F(3,32)=1.43; P=0.25
Probability of species detection	Global	F(5,93)=4.80***	F(3,93)=94.28***	F(15,93)=1.48; P=0.13
	Pore sizes	F(2,46)=1.89; P=0.16	F(3,46)=48.79***	F(6,46)=1.13; P=0.36
	Filter types	F(1,31)=4.90*	F(3,31)=28.27***	F(3,31)=2.39; P=0.09
	Pre-filtration	F(1,32)=0.65; P=0.43	F(3,32)=32.54***	F(3,32)=2.85; P=0.05
Horn index	Global	F(5,204)=14.09***	F(3,204)=34.67***	F(15,204)=6.55***
	Pore sizes	F(2,100)=10.33***	F(3,100)=30.29***	F(6,100)=9.31***
	Filter types	F(1,68)=53.63***	F(3,68)=5.18**	F(3,68)=4.29**
	Pre-filtration	F(1,72)=34.96***	F(3,72)=24.86***	F(3,72)=24.29**
Correlation with abundance	Global	F(5,93)=1.58; P=0.17	F(3,93)=4.48*	F(15,93)=1.05; P=0.41
	Pore sizes	F(2,46)=3.22*	F(3,46)=3.73*	F(6,46)=1.94; P=0.09
	Filter types	F(1,31)=0.05; P=0.83	F(3,31)=1.70; P=0.19	F(3,31)=0.58; P=0.63
	Pre-filtration	F(1,32)=0.0025; P=0.96	F(3,32)=5.79**	F(3,32)=0.69; P=0.56
Correlation with biomass	Global	F(5,93)=2.30; P=0.051	F(3,93)=8.85***	F(15,93)=1.51; P=0.11
	Pore sizes	F(2,46)=5.80**	F(3,46)=12.31***	F(6,46)=2.61*
	Filter types	F(1,31)=0.005; P=0.95	F(3,31)=2.93*	F(3,31)=0.81; P=0.50
	Pre-filtration	F(1,32)=0.44; P=0.51	F(3,32)=7.53***	F(3,32)=0.21; P=0.89

Note: The compared treatments in three different groups are: pore sizes (0.45MCE, 0.8MCE and 1.2MCE), filter types (0.45MCE and 0.45Sterivex) and pre-filtration (0.45MCE and PF_0.45MCE). Replicates identified as outliers are excluded. Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.