Environmental nucleic acids: a field-based comparison for monitoring freshwater habitats using eDNA and eRNA

Running title: Comparing eDNA and eRNA for biomonitoring

Joanne E. Littlefair*[1, 2], Michael D. Rennie[3,4], Melania E. Cristescu[1]

* Corresponding author j.e.littlefair@qmul.ac.uk

Author affiliations:

[1] Department of Biology, McGill University, 1205 Docteur Penfield, Montréal, QC, Canada

[2] Current affiliation: Queen Mary University of London, Mile End Road, London, UK

[3] IISD Experimental Lakes Area, 111 Lombard Ave., Suite 325, Winnipeg, MB, Canada

[4] Department of Biology, Lakehead University, 955 Oliver Road, Thunder Bay, ON, Canada

Abstract

Nucleic acids released by organisms and isolated from environmental substrates are increasingly being used for molecular biomonitoring. While environmental DNA (eDNA) has received attention recently, the potential of environmental RNA as a biomonitoring tool remains less explored. Several recent studies using paired DNA and RNA metabarcoding of bulk samples suggest that RNA might better reflect “metabolically active” parts of the community. However, such studies mainly capture organismal eDNA and eRNA. For larger eukaryotes, isolation of extra-organismal RNA will be important, but viability needs to be examined in a field-based setting. In this study we evaluate (a) whether extra-organismal eRNA release from macroeukaryotes can be detected given its supposedly rapid degradation, and (b) if the same field collection methods for eDNA can be applied to eRNA. We collected eDNA and eRNA from water in lakes where fish community composition is well documented, enabling a comparison between the two nucleic acids in two different seasons with monitoring using conventional methods. We found that eRNA is released from macroeukaryotes and can be filtered from water and metabarcoded in a similar manner as eDNA to reliably provide species composition information. eRNA had a small but significantly greater true positive rate than eDNA, indicating that it correctly detects more species known to exist in the lakes. Given relatively small differences between the two molecules in describing fish community composition, we conclude that if eRNA provides significant advantages in terms of lability, it is a strong candidate to add to the suite of molecular monitoring tools.

Keywords: environmental DNA, environmental RNA, metabarcoding, freshwater, fish, seasonality

Introduction

Environmental nucleic acids (eNAs) such as environmental DNA (eDNA) and environmental RNA (eRNA) are emerging as reliable methods for monitoring aquatic biodiversity (Cristescu & Hebert, 2018; Deiner et al., 2017). One concern with the recovery of eDNA revolves around the dynamics of stability and persistence of nucleic acids in the environment that can lead to false detections of local species. There are several possible scenarios through which false positives might result from eDNA; for example, the transport of eDNA molecules from an upstream to a downstream location, or the resuspension of eDNA from a sediment layer which originates from older communities (Corinaldesi, Beolchini, & Dell'Anno, 2008) to the water column. Situations in which it might be difficult to distinguish transported/residual eDNA signal from true signal include detections at low abundance from rare species and at invasion fronts (Jerde, Mahon, Chadderton, & Lodge, 2011). The degree of transported signal relies on the complex interplay among abiotic factors influencing the release, degradation, and persistence of eNAs, the speed and volume of substrate flow, and biotic factors such as biomass, metabolic rates, and behaviour which determine the volume of eNAs released from animal populations. For this reason, it is difficult to predict the extent of transported eNA in each individual scenario. Laboratory and field studies based on eDNA detect signal across distances from several metres to tens of kilometres (Deiner & Altermatt, 2014; Deiner, Fronhofer, Mächler, & Altermatt, 2016; Jane et al., 2015; Jerde et al., 2016; Shogren et al., 2017). This key property of eDNA sampling has important consequences for separating the presence of active communities with molecular monitoring from residual signal.

Recently, it has been proposed that eRNA could be more labile than eDNA and is therefore a candidate molecule for reducing problems associated with transported signal (Cristescu, 2019). The increased lability of eRNA when compared with eDNA is thought to originate from its single-stranded structure, the presence of additional hydroxyl groups allowing for

base catalysed hydrolysis (Y. Li & Breaker, 1999), and the ubiquitous presence of exogenous and endogenous RNases (Tan & Yiap, 2009). These characteristics are thought to lead to a faster rate of degradation of eRNA when compared with eDNA; for example, eRNA has a 4-5 hour faster half-life when compared with eDNA (Marshall, Vanderploeg, & Chaganti, 2021). Thus, it may be possible for eRNA signal to distinguish biologically active communities from dead/dormant ones, and local communities from transported molecular signal when eRNA is applied to species monitoring (Barnes & Turner, 2016; Deiner et al., 2017; Pawlowski et al., 2018). For example, eDNA in ballast water was found to contain 57% OTUs assigned to fungi which are thought to represent legacy OTUs, whereas OTUs detected by eRNA included mainly active metazoa and ciliates. Pawlowski et al. (2014) found that DNA recovered greater benthic taxonomic richness when compared with RNA, which could be explained by the detection of previous benthic successions of DNA, as opposed to solely cellularly active taxa (see also Guardiola et al., 2016 for similar findings). The same study found that RNA detected benthic community responses to fish farming to a greater degree than DNA (Pawlowski et al., 2014). Similarly, Dowle et al (2015) found moderately stronger correlations between bacterial RNA OTU data and environmental indices, when compared with DNA OTUs. Although these studies are suggestive, eRNA has received much less attention than eDNA to date, particularly with respect to molecular monitoring of macro-eukaryotes.

Before eRNA is used more commonly in biomonitoring applications, its efficiency of recovery and accuracy of representation of known biological communities must be assessed. Although studies based on organismal RNA are valuable, it is important to examine the viability of extra-organismal RNA recovery under natural field conditions, given that factors such as temperature, acidity, and microbial activity have already been shown to influence eDNA persistence (Sansom & Sassoubre, 2017; Seymour et al., 2018; Strickler, Fremier, & Goldberg, 2015; Tsuji, Ushio, Sakurai, Minamoto, & Yamanaka, 2017). There are many

77 practical considerations when applying molecular monitoring methods in the field, such as
78 the time between sample collection and storage, and choice of buffers used to preserve the
79 molecules until lab work can be undertaken (Deiner, Walser, Mächler, & Altermatt, 2015;
80 Dickie et al., 2018). If eRNA cannot be used in a "field" scenario in the same way that eDNA
81 is, any theoretical benefits due to increased molecule lability would be outweighed by
82 practical disadvantages. Moreover, eRNA may be too labile to be reliably sampled in the
83 field in the same way that eDNA is. The few studies that provide comparisons of community
84 recovery with DNA and RNA metabarcoding focussed on metabarcoding of bulk samples
85 (organismal RNA) which primarily includes smaller eukaryotes and bacteria in marine
86 sediment samples (Brandt et al., 2020; Guardiola et al., 2016; Laroche et al., 2018, 2017;
87 Orsi, Biddle, & Edgcomb, 2013; Pawlowski et al., 2016). Larger animals such as vertebrates
88 (e.g. fish and mammals) might be detected by considering extra-organismal molecules
89 isolated from the environment which are not the primary focus of bulk sample studies. Few
90 studies have specifically considered extra-organismal eRNA, although eRNA was compared
91 to eDNA using ddPCR from two species in laboratory aquaria (Wood et al., 2020), and
92 Miyata et al., (2021) recently used eRNA metabarcoding to detect fishes from two sites in
93 the Naka river, Japan.

94 In our study, we sampled and filtered water from multiple freshwater lakes, extracting eDNA
95 and eRNA from paired filter halves to compare the community composition of fishes
96 recovered by each molecule. We targeted extra-organismal nucleic acids (eDNA and eRNA)
97 from fish by generating amplicon libraries using the MiFish-U 12S rRNA fragment (Miya et
98 al., 2015). We compared the sampling effort required, species detection success and
99 recovered community composition of these two molecules, and compared both against long-
100 term biomonitoring data collected by conventional methods at these same sites. Sampling
101 was conducted in the summer and again in autumn to capture distinct differences in lake
102 thermal structure (stratification during summer and turnover during autumn). We expected

the recovered molecular-based community composition to vary between these two sampling periods, due to differences in water stratification and mixing regimes which control how eDNA is distributed in the lakes, and according to the changing seasonal depth preferences of the fishes (Littlefair, Hrenchuk, Blanchfield, Rennie, & Cristescu, 2021).

Methods

Field collection

Samples of eDNA and eRNA were collected in parallel to maximize comparability. We sampled a total of seven lakes (two in 2017 and five in 2018; Table S1) at the IISD Experimental Lakes Area (IISD-ELA), a freshwater research facility in northwest Ontario, Canada. Biomonitoring has been ongoing at IISD-ELA since the 1960s, providing a well-developed knowledge of species composition in this area and in these lakes specifically (Table S2).

Six 500ml water samples were collected from each lake along a depth gradient at the deepest point of each lake, by dividing the water column into six even parts and using an electrical pump and PVC tubing secured to a weight to collect the water at the relevant depth. Collecting on a depth gradient maximises the chance of recovering cold-water fish species such as lake trout (*Salvelinus namaycush*) (Littlefair et al., 2021). Water samples were pumped into sterile Whirl-Pak bags and sealed within an individual large Ziplock bag. Tubing was cleaned between each depth point by pumping one litre of 30% bleach, one litre of distilled water, followed by lake water at the new sampling depth for a two-minute period prior to sample collection. Separate tubing was used for each lake. All samples were immediately transported to the field laboratory in a cooler with ice packs and stored at 4 ºC until filtration. Filtering of samples took place within 2-5 hours of sample collection. Water samples were filtered onto 47 mm 0.7 µm pore GF/F filters using an electric vacuum pump and filtering manifold (Pall Corporation, ON, Canada) in a room which was not used for

128 animal work. Each filter was cut into two using an individual pair of forceps and scissors. Half
129 of the filter was immediately frozen at -20°C for DNA analysis and half was preserved in
130 370µl RLT buffer (Qiagen) with 1% β-mercaptoethanol and then frozen at -20°C for RNA
131 analysis. One negative control of 500 ml distilled water was stored in the cooler and filtered
132 in the same way as the field samples for each lake. In total, 84 eDNA and 84 eRNA samples
133 were taken across the entire study. Filters were shipped on dry ice to McGill University,
134 Montréal for molecular analysis.

135 **Molecular analysis**

136 Extractions of eRNA were performed from the first half of the filter using the Qiagen RNEasy
137 Mini kit with some modifications to accommodate the filter. Filters were vortexed for 20
138 seconds and centrifuged in the RLT/ β-mercaptoethanol buffer for 3 minutes at 14,000rpm. A
139 total of 325µl of this buffer was mixed with 325µl ethanol and the rest of the procedure
140 followed the kit protocol intended for extracting total RNA from animal cells. The eRNA was
141 resuspended in two elutions of 30µl RNAse free water to give a final volume of 60µl.

142 Extractions of eDNA from the second half of the filter were performed using the Qiagen
143 Blood and Tissue kit. We followed the manufacturer's instructions except that 370µl of ATL
144 buffer was used in an initial overnight incubation step. After extraction, both eDNA and
145 eRNA were preserved at -80°C.

146 To avoid DNA contamination in the eRNA samples, DNA was digested from 20µl eRNA
147 extracts with the DNA-free™ DNA Removal Kit (ThermoFisher Scientific) following the
148 manufacturer's instructions and using 2µl DNase I Buffer, 1µl rDNase I and 2µl DNase
149 Inactivation reagent. Samples were checked for residual contaminating DNA using PCR
150 amplification using the MiFish-U primers tagged with Illumina sequencing adapters (Miya et
151 al., 2015). These primers target a hypervariable region of the 12S rRNA locus (163-185bp in
152 length) which has previously been used to characterise the fish community in this area and

provides good species level discrimination of ASVs (Littlefair et al., 2021). We used the following PCR chemistry: 7.4µl nuclease free water (Qiagen), 1.25µl 10X buffer (Genscript), 1 mM MgCl2 (ThermoFisher Scientific), 0.2mM GeneDirex dNTPs, 0.05mg bovine serum albumen (ThermoFisher Scientific), 0.25mM each primer, 1U taq (Genscript) and 2µl DNA in a final volume of 12.5µl. We followed a touchdown thermocycling protocol which we have found reduces the amount of non-specific amplification (bacterial taxa) at this locus: 95°C for 3 minutes, 12 cycles of touchdown PCR (98°C for 20 seconds, 66°C for 15 seconds decreasing by 0.2°C each time, 72°C for 15 seconds) followed by 28 cycles with an annealing temperature of 64°C, 72°C extension for 5 minutes. No residual contaminating DNA was found in eRNA samples. A total of 10µl of sample was therefore reverse transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (ThermoFisher Scientific) in 20µl reactions following the kit instructions.

We then amplified the cDNA and eDNA in triplicate 12.5µl reactions following the MiFish-U PCR protocol laid out above, and checked amplification using 1% agarose gels with SYBR Safe. We then combined the triplicate reactions into one sample and performed a cleanup with AMPure beads. Cleaned amplicons were then dual-indexed using the Nextera v2 DNA indexes, cleaned again, equimolarised to 3ng/µl and sent for sequencing at Génome Québec, Montréal. Samples were sequenced using 2 x 250bp paired end sequencing with an Illumina MiSeq.

To prevent contamination, we processed the samples in a clean, pre-PCR dedicated lab. Before beginning any eRNA work we thoroughly cleaned benches with 10% bleach solution and RNAse wiper. Laboratory equipment was cleaned with 70% ethanol and RNAse wiper. Negative controls were included at each major step: field sampling, RNA/DNA extraction, reverse transcription, and PCR amplification.

Bioinformatics

We used custom scripts to remove adapters, merge paired sequences, check quality and generate amplicon sequencing variants (ASVs). Samples were received as demultiplexed fastq files from Génome Québec. Non-biological nucleotides were removed (primers, indices and adapters) using cutadapt (Martin, 2011). Paired reads were merged using PEAR (Zhang, Kobert, Flouri, & Stamatakis, 2014). Quality scores for sequences were analysed with FASTQC (Andrews, 2010). Reads were length filtered between 152-192 bp. Amplicon sequencing variants (ASVs) were generated using the UNOISE3 package (Edgar, 2016), which uses a denoising pipeline to remove sequencing error and to cluster sequences into single variants (100% similarity). The full bioinformatics pipeline is available from https://github.com/CristescuLab/YAAP. After ASVs were generated, we assigned taxonomy using BLAST+ (Camacho et al., 2009) and BASTA (Kahlke & Ralph, 2019), a last common ancestor algorithm. We used a custom reference database which contained only fish known to exist in the Lake of the Woods region (Ontario, CA), downloaded from the NCBI database on 12 August 2018. We also compared our assignments against the full NCBI database and found only one additional fish ASV with the larger database. Other taxonomic groups appeared at very low frequencies when our ASVs were matched against the NCBI database, such as bacterial, mammalian and bird taxa.

Statistical analysis

All analyses were conducted in R v4.0.2.

Bioinformatic filtering

Differences in the final library sizes of eDNA and eRNA filter halves after bioinformatic filtering were analysed by performing a paired *t*-test. We also explored the correlation between library sizes in paired filter halves with a Spearman's rank correlation test.

Sampling effort

For these analyses, ASV count data was converted to incidence data. The number of water samples required to adequately sample species richness was assessed by creating sample accumulation curves using the function specaccum in the "vegan" package (Oksanen et al., 2019). Dataframes were filtered to only include fish species (i.e. the small amount of non-target taxa were removed) in order to draw comparisons with conventional fishing techniques for surveying biodiversity. The differences between eDNA and eRNA accumulation curves were assessed by plotting separate curves for each molecule within each lake and season. Species richness according to conventional techniques was plotted on the graphs with a grey dashed line. We also plotted species accumulation curves with the full dataset (i.e. with non-target taxa included); these are presented in the supplementary material. The ability of each molecule to achieve adequate sampling of species richness was determined by calculating observed sample coverage for each molecule, season and lake combination using iNEXT. We also looked at how species richness varied with increasing sample coverage for each molecule, season and lake combination.

Species detection

The fish species composition of the lakes is well known as a result of decades of ongoing monitoring. It was therefore possible to assess the relative performance of eDNA and eRNA to determine species composition against conventional techniques. Fish species were recorded as being present according to conventional techniques if they were consistently detected by typical collection methods (trap netting and short-set gill netting), as part of an ongoing broad scale monitoring program (using sampling procedures as outlined in Rennie et al., 2019, which are typical of sampling efforts in the IISD-ELA lakes in the current study). Using these techniques, we recorded 104 detections with conventional techniques across all the lakes and species. A species was recorded as being present in a lake according to molecular methods if it was detected in at least one of the six water samples taken from that

lake. The number of true positives detected by eDNA and eRNA was expressed as a fraction of the total number of conventional detections possible across all lakes and seasons.

We then calculated the true positive rate as the number of detections made with molecular methods as a proportion of true positives and false negatives (which we defined according to the results of conventional sampling). We additionally calculated the false discovery rate as the number of false positives (i.e. species we know are not present in the lakes according to conventional sampling) as a proportion of true positives and false positives (i.e. all detections). True positive and false discovery rates are positive numbers on the scale of 0 – 1, with a higher number indicating a larger proportion of true positives or false discovery in the data. We then used a mixed effects model (fitted with glmmTMB) to examine whether the true positive rate and the false discovery rates differed significantly between eDNA and eRNA, using sampling season as a covariate and fitting the lake as a random effect. We tested all models for overdispersion and examined model residuals using Dharma (Hartig, 2021), and tested the significance of each explanatory term by fitting nested models using the “drop1” command with a chi-squared distribution.

Community composition

nMDS was applied to a Bray-Curtis dissimilarity matrix to visually explore the differences in community composition between eDNA and eRNA for each lake. 95% confidence ellipses were drawn around each season/molecule grouping using the ordiplot and ordiellipse functions in vegan. We used the manyglm function in the mvabund package to fit multispecies GLMs for hypothesis testing (Wang, Naumann, Wright, & Warton, 2012). We tested the effects of environmental nucleic acid type (eDNA/eRNA) and season (August/October) as predictors on the community dataset, as well as the interaction between these two factors. When we tested the interaction term, it was not statistically significant, so we removed it and fitted a new additive model with main effects only. We included library

size as a log offset to account for library size variation between samples. We tested both poisson and negative binomial distributions and found that the negative binomial distribution removed patterns in the model residuals, so we retained this distribution for our models. We used the anova.manyglm function to retrieve test statistics using adjusted *p*-values to account for multiple testing (i.e. the detection of multiple species). We accounted for the block design of sampling multiple lakes by restricting permutations to within-lake blocks, by supplying a permutation matrix to the bootID argument in the anova function designed using the permute function in vegan.

Results

Bioinformatic filtering

Initial library sizes before bioinformatic filtering for eRNA were on average 6.22% smaller than for eDNA. Similar amounts of sequences were removed for both molecule types during the process of adapter removal, pair merging and final trim of primers (Table 1). However, many more sequences were removed during the length filtering step for eRNA when compared with eDNA (eRNA = 25.4% removed, compared with eDNA = 2.43% removed), indicating that the amplification and sequencing of eRNA resulted in more sequences outside the 152-192bp length filter. Despite significantly smaller average library sizes for eRNA compared with eDNA after bioinformatic filtering (paired *t*-test, $t = 4.09$, df = 97, $p < 0.001$), the denoising steps produced similar amounts of ASVs for both molecules (eDNA = 107, eRNA = 115), and similar percentages of sequences mapped onto these ASVs (eDNA = 98.5%, eRNA = 98.4%). There was a moderate but significant correlation in sequence numbers of the filtered library size from paired eDNA and eRNA extracted from the same water sample (Spearman's rho = 0.539, $p < 0.001$).

Sampling effort

Species accumulation curves were more rapid for eRNA taken in October than for eDNA or eRNA in August in 4/7 lakes, but were otherwise inconsistent among lakes, within or across seasons (Figure 1). Generally only a very small number of water samples (three - four) per lake were needed in order to achieve a plateaued species accumulation curve. Compared with conventional techniques, molecular techniques sometimes under- or over-sampled fish species richness, but generally only by one or two species (see below). When considering the entire dataset (i.e. fish and non-target taxa), the curves showed a higher species richness than expected in the lakes based on conventional sampling (Figure S1), primarily because the MiFish-U marker detects small numbers of taxonomic groups other than fish (e.g. zooplankton, human DNA, birds; see below). Mean sample coverage for eDNA was 0.882 and for eRNA was 0.863 (Table S3). Although we did not sample sufficient lakes to conduct a formal statistical analysis on the differences in sample coverage between eDNA and eRNA, there were no noticeable visual trends (Figure S2).

Species detection

The vast majority of sequences (95.8%) in this dataset were assigned to Actinopterygii ASVs. A total of 15 Actinopterygii ASVs were detected by both eDNA and eRNA, with an additional four detected by eRNA only and an additional six detected by eDNA only. Non-target ASVs were also generated, although only a relatively small percentage of sequences in the dataset were actually assigned to these ASVs (4.23%). Non-target ASVs varied between the two molecules: eRNA detected a greater incidence of bacterial, algal and arthropod ASVs, and eDNA detected more mammalian and unassigned ASVs.

When using the database consisting of fish found in the northwest Ontario region, ASVs could generally be assigned at species level. There were two fish species which were not detected by either molecule (*Culaea inconstans* and *Rhinichthys cataractae*). Moreover, *Chrosomus neogaeus* and *Chrosomus eos* could only be detected at genus level from a

single ASV: species specific identification is not possible due to the presence of mitochondrial hybrids in the region (Mee & Taylor, 2012). We therefore counted detections only once for the *Chrosomus* genus to avoid double counting.

There were very small amounts of sequences detected in negative controls after bioinformatic filtering. These controls had an average library size of 178, compared with eDNA/eRNA libraries which had an average size of 69,408. Of these, 91.7% of sequences in the negative controls matched fish from northwest Ontario (rather than other taxa such as bacteria or mammals). In almost all cases, the sequences matched the species composition from the lake that the negative control originated from, indicating that contamination of the negative control originated from within the lake sampled, rather than across-lake contamination or tag jumping. We did not find any amplification in the eRNA after the use of the DNA removal kit before conversion to cDNA.

Sampled eRNA had a small but significantly greater true positive rate than eDNA, indicating that eRNA correctly detected more of the species known to exist in the lakes based on conventional sampling (Figure 2A, eRNA true positive rate: 0.692 per sample, eDNA true positive rate: 0.648 per sample, $p = 0.0043$). There was no difference in false discovery rate between the two molecules; i.e., neither molecule detected more false positives as a proportion of all detections (Figure 2B, eRNA: 0.052, eDNA: 0.046, $p = 0.568$). There was a significantly lower false discovery rate in August samples when compared with those collected in October (August: 0.030, October: 0.069, $p = 0.0004$). Usually false positive detections were of a low read count (Table S2).

Community composition

Using NMDS as a visual technique to explore the differences in community composition as explained by different predictors, the largest differences in community composition were generated by the differences in species composition between lakes (Figure S3). Samples

collected in October were more similar to each other in terms of community composition than samples collected in August (Figure S4), although these two distributions were nested within each other. In the overall dataset, eDNA and eRNA samples detected largely similar communities, although eRNA samples were slightly more similar to each other than eDNA samples (Figure S5). Within a single lake, there was variation in whether eDNA and eRNA detected similar community compositions (Figure 3). Although consistent differences between the two molecules were not evident, there did seem to be a stronger seasonal effect, as samples that were collected in August were often more dissimilar to each other than samples collected in October. This is reflected in a larger and partially non-overlapping 95% confidence ellipse for these groups of samples.

We visually explored the proportion of sequences in each eDNA and eRNA sample which belonged to each species or taxa in August (Figure 4A) and October (Figure 4B). We calculated the proportion of sequences belonging to each species out of the total library size per filter. In some taxa, eDNA and eRNA samples had strikingly similar proportions of sequences within each library, for example white sucker (*Catostomus commersonii*), finescale/northern redbelly dace (*Chrosomus spp.*), slimy sculpin (*Cottus cognatus*), and fathead minnow (*Pimephales promelas*) (see also Table S2). In a few species, there were differences between the proportions of nucleic acids; for example, there was always a higher proportion of *Coregonus artedi* eDNA sequences compared to eRNA, and always a lower proportion of *Perca flavescens* eDNA sequences compared to eRNA sequences. There were also seasonal differences in the proportion of sequences belonging to each species; for example, there were more sequences belonging to lake trout (*Salvelinus namaycush*) in October compared with August, which reflects the spawning times and patterns of habitat occupancy for this fish (Littlefair et al., 2021). Similarly, there were also seasonal effects on sequence numbers of *Coregonus artedi* and *Cottus cognatus* which are both cold water fish (higher concentrations in October), and much lower concentrations of *Perca flavescens*

sequences in October. These visual differences were reflected by the results of the multispecies GLM, which retained significant effects for molecule type (df = 1, 166, deviance = 284.7, p = 0.02) and season (df = 1, 165, deviance = 296.4, p = 0.005). There was no significant effect of the interaction between molecule type and season on the numbers of sequences, indicating that eDNA and eRNA detection of taxa did not respond differently between the two seasons.

Discussion

Traditionally, RNA has been thought of as a very labile molecule, too difficult to extract and preserve in a field setting. However, we have shown here that eRNA achieves similar rates of macroeukaryotic species detection when compared with eDNA within the context of sampling for environmental assessment, and in fact had a slightly higher rate of true positive detection per sample than eDNA. The majority of RNA studies to date have focused on metabarcoding of bulk samples (organismal RNA), but the choice of fish as a study organism shows that it is possible to detect animals based simply on extra-organismal RNA released into the environment. This is the first paper to solely focus on the release of extra-organismal RNA in lakes with comparisons to both eDNA and well-documented conventional monitoring techniques, and these comparisons support its application more broadly to assess species presence/absence and ecosystem functioning.

Extra-organismal RNA is stable enough to collect with comparable field techniques used with extra-organismal DNA. Across the entire study, all target fish species which were detected with eDNA could also be detected with eRNA. Per sample, the true positive rate of detection was significantly higher with eRNA than eDNA, reflecting the results of Miyata et al., (2021), although the difference in detection rate between the two molecules was small. This is consistent with studies based on bulk DNA and RNA which report substantial overlap between OTU or species detections by the two molecules (e.g. Pochon, Zaiko, Fletcher,

Laroche, & Wood, 2017). In our study, field collection involved pumping water from the lake, transporting the water samples back to the field laboratory, and filtering them, which took place between two and five hours after collection, a protocol typical of many eDNA studies (Balasingham, Walter, Mandrak, & Heath, 2018; Bylemans, Gleeson, Hardy, & Furlan, 2018; Hänfling et al., 2016; Jeunen et al., 2019; J. Li, Lawson Handley, Read, & Hänfling, 2018; Stat et al., 2018; Zhang et al., 2020). Here, we have shown that extra-organismal eRNA can withstand comparable field methods and can be filtered and sequenced from the water column, performing comparably to eDNA for species detections when compared with conventional monitoring techniques.

While false negatives occurred for both eDNA and eRNA in certain lakes when molecular results were compared to the known species composition of the lake ascertained by conventional techniques, the incidence of these false negatives seemed to be linked to the ecology of these species rather than the type of nucleic acid molecule. For example, fish species which could not be detected in certain lakes were almost always recorded by the biomonitoring program as being at rare or moderate levels of abundance in that particular lake (e.g., *Couesius plumbeus* in lake 626). Moreover, species which favoured a littoral habitat or which live in small inlets around the lakes (e.g., *Chrosomus* spp, *Catostomus commersoni,* and *Esox lucius*) were also recorded as false negatives in some lakes. This is perhaps not surprising given that sampling took place at the centre point of the lakes, well away from the shoreline. The two fish which were reported as being present by conventional techniques but not detected by molecular sampling in any lake habitat (*Culaea inconstans* and *Rhinichthys cataractae*) were likewise recorded as being at “rare” levels of abundance and were also species which favoured littoral or inlet habitats around the periphery of the lakes. For example, only one *Rhinichthys cataractae* was caught in the survey in 2017, and none have been caught in subsequent years.

There was no significant difference in the false discovery rate between eDNA and eRNA. False discovery rate is defined as the number of false positives as a proportion of the sum of true positives and false positives. Some studies have proposed that eRNA might detect cellularly active taxa only, as opposed to dead and dormant taxa or resuspended sedimentary DNA, and thus minimise the false discovery rate when compared with eDNA (Dowle et al., 2015; Pawlowski et al., 2014; Pochon et al., 2017; Visco et al., 2015, although see Brandt et al., 2020), although to date laboratory degradation experiments indicate that eRNA might not degrade significantly faster than eDNA (Wood et al., 2020). Further studies in a field setting will be needed to determine the advantages of eRNA of overcoming false positives detected by eDNA sampling, should they exist. In semi-natural settings, experiments have been performed using caged animals or artificial spikes of DNA to assess the effects of time or distance on the degradation of the DNA signal (Harper, Anucha, Turnbull, Bean, & Leaver, 2018; Jane et al., 2015; Pilliod, Goldberg, Arkle, & Waits, 2014). Observational field studies have also provided evidence that DNA flows downstream from populations (Deiner & Altermatt, 2014; Deiner et al., 2016). These methodologies could be performed in parallel with eRNA to analyse whether the use of eRNA improves the false positives currently detected by eDNA. Some of the lakes that we examined in this study are headwater lakes (Table S1), meaning that there is no opportunity for water and nucleic acids to flow from upstream populations, so the lack of difference in false discovery rate between eDNA and eRNA is perhaps unsurprising in this context.

We did not specifically analyse the degradation time of eRNA in our study, but it is noteworthy that a reliable signal from this nucleic acid could still be detected two to five hours after collection in the field and storage in refrigerated conditions. This and several other lines of evidence point to extra-organismal RNA being more stable than previously thought (Cristescu, 2019; Torti, Lever, & Jørgensen, 2015). It may be that the eRNA is temporarily stable inside cells, organelles or vesicles; a study of long noncoding RNA

(lncRNA) found that its half-life within cells ranged from <2 to >16 hours (Clark et al., 2012). Although Marshall et al., (2021) found that eRNA had a 4-5hour faster half-life than eDNA, the half-life was still within 8.84 – 13.54 hours (depending on marker selection and RNA type), which was within the collection window of our study. Alternatively, eRNA could be combined with organic or inorganic particles within the water column which aids with stability as with eDNA; for example, Wood et al., (2020) found that eRNA could still be detected in biofilms from 4/15 aquaria at the end of a 21-day experiment. There have also been suggestions that RNA might preserved for long time periods in sediments by binding to sedimentary particles (Cristescu, 2019; Orsi et al., 2013; Torti et al., 2015). For eRNA to be a successful complement to eDNA to resolve spatial-temporal issues in species detection it must be stable enough to be detected, but ideally degrade faster than eDNA, in order to provide a detection signal which is in closer geographic proximity to the population in question.

We found strong effects of season and molecule type on the proportion of sequences assigned to different fish species in the libraries. In many (but not all) cases, this seems to correspond to aspects of species biology and abundance in the lakes. We found higher proportions of sequences for both molecule types in October for cold water species such as lake herring (*Coregonus artedi*), slimy sculpin (*Cottus cognatus*), and lake trout (*Salvelinus namaycush*) (Figure 4B). The October sampling corresponded to the spawning time of lake trout, an event which has been linked to the creation of eNAs (Tillotson et al., 2018), and although the other two species were not spawning during October, their activity expands from deep, hypolimnetic waters to the entire lake as lake surface temperatures fall below thermal maxima and towards optimum temperatures for these species (Hasnain, Escobar, & Shuter, 2018). Across both seasons, there were high levels of yellow perch (*Perca flavescens*) sequences, with some libraries containing >90% reads assigned to yellow perch. This may reflect the greater abundance of yellow perch in these lakes relative to other

454 species. The numbers of yellow perch sequences in August were particularly high; this may
455 be due to the relative inactivity of other species in warmer parts of the lake during the
456 summer, as the proportion of sequences from one species affects the proportion of
457 sequences from others in the sequencing libraries. There were also consistent effects of
458 molecule type on the proportions of sequences which was largely consistent across
459 seasons; for example, high levels of lake herring eDNA relative to eRNA, and high levels of
460 yellow perch eRNA relative to eDNA in both August and October. However, for many
461 species, the consistency of the relative levels of eDNA and eRNA was surprisingly strong,
462 across both seasons. The reasons for these patterns are, as yet, unknown.

463 Comparing workflows will be important when assessing the relative strengths of these two
464 molecules. Surprisingly, eRNA was robust to a typical eDNA field protocol, which involved a
465 time lag of 2-5 hours between collection and the filtering and storage of filters (typical of
466 many eDNA studies). Possible increases in yield could be achieved by adapting or applying
467 recent innovations which involve the capture and storage of molecules *in situ* or direct
468 sequencing in the field (Truelove, Andruszkiewicz, & Block, 2019). We applied equivalent
469 levels of care to preserving the stability of both molecules after collection; for example,
470 keeping filters frozen at all times, shipping on dry ice, use of diluted bleach and RNase in the
471 laboratory to clean surfaces and equipment, and working with samples on ice. Inherently, the
472 molecular workflow involves some differences between the two molecules. We chose to use
473 spin column based kits for extractions (Blood and Tissue kit for DNA, RNeasy kit for RNA,
474 both manufactured by Qiagen), as these methods have been shown to yield some of the
475 highest quantities of RNA (Tavares, Alves, Ferreira, & Santos, 2011). Alternative options are
476 provided by kits which co-extract the two molecules together such as the ZR-Duet DNA/RNA
477 MiniPrep kit from Zymo Research (as used in Pochon, Zaiko, Fletcher, Laroche, & Wood,
478 2017). Sampled eRNA then requires additional steps to convert the molecule into cDNA for
479 sequencing, involving DNA digestion and reverse transcription, before performing an

equivalent PCR amplification as in traditional DNA metabarcoding. These additional steps might involve the loss of molecule yield. We found that the final recovered sequence counts were lower for eRNA libraries – given that an equimolar amount was added to the metabarcoding libraries, this loss of molecules possibly occurred through the removal of low-quality reads during the bioinformatics pipeline. As a final note, these additional molecular steps mean that the extraction and processing of eRNA is more costly per sample in terms of kits and personnel time in comparison to eDNA, which may be an important consideration currently when deciding the relative benefits between eDNA and eRNA in field studies.

However, the development of newer sequencing technologies mean that the simplification of this workflow is on the horizon. Future sequencing technologies will mean that RNA can be sequenced directly without conversion to cDNA or with PCR bias which results from amplification steps. The possibility of starting with low-input amounts means that this might be particularly suitable for eRNA applications, and this technology has already been applied to microbial mock communities (Nicholls, Quick, Tang, & Loman, 2019). Some features which will be of interest to eDNA/eRNA scientists are still in development at the time of writing, such as the ability to multiplex samples with direct RNA sequencing kits, but further advances will be anticipated with interest.

Acknowledgements

We are indebted to many IISD Experimental Lakes Area students and staff for maintaining records of field data and for logistical assistance with this project. S Michaleski, R Henderson, P Bulloch, M Haust, C Jackson and A McLeod contributed specific field assistance to this project, and L Hrenchuk oversaw fieldwork. K Sandilands provided the pumping equipment used in this project. Dr JS Hleap provided bioinformatics support to this project. M Harris provided training on the eRNA extraction and reverse transcription protocols used in this work. This work was funded by a Mitacs Accelerate Industrial

Fellowship (JEL), an NSERC Collaborative Research and Development award (MEC), Canada Research Chair and NSERC Discovery awards to MEC and MDR, Québec Centre for Biodiversity Science Excellence award (JEL), the WSP Montréal Environment Department and in-kind support from the IISD Experimental Lakes Area and Fisheries & Oceans Canada. We thank Jean Carreau and Patrick LaFrance of WSP Montréal for useful discussions on the topics of eDNA and biomonitoring.

Animal permits

Scientific collection permits to facilitate fish collections were authorised through the Ontario Ministry of Natural Resources and Forestry. Handling of fish was carried out under the authority of the Government of Canada through Fisheries and Oceans Canada (prior to and including 2013), and with the approval of Animal Care Committees through the University of Manitoba (2014; permit #F14-007) and Lakehead University (2015–present, permit #s 1464656 and 1466997).

Word count: 6576 without references

References

Balasingham, K. D., Walter, R. P., Mandrak, N. E., & Heath, D. D. (2018). Environmental DNA detection of rare and invasive fish species in two Great Lakes tributaries. *Molecular Ecology*, *27*(1), 112–127. https://doi.org/10.1111/mec.14395

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

Brandt, M. I., Trouche, B., Henry, N., Liautard-Haag, C., Maignien, L., de Vargas, C., … Arnaud-Haond, S. (2020). An Assessment of Environmental Metabarcoding Protocols Aiming at Favoring Contemporary Biodiversity in Inventories of Deep-Sea Communities. *Frontiers in Marine Science*, *7*, 234. https://doi.org/10.3389/fmars.2020.00234

Bylemans, J., Gleeson, D. M., Hardy, C. M., & 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*, 8697–8712. https://doi.org/10.1002/ece3.4387

Clark, M. B., Johnston, R. L., Inostroza-Ponta, M., Fox, A. H., Fortini, E., Moscato, P., … Mattick, J. S. (2012). Genome-wide analysis of long noncoding RNA turnover. *Genome Research*, *22*, 885–898. https://doi.org/10.1007/978-1-4939-2253-6_19

Corinaldesi, C., Beolchini, F., & Dell'Anno, A. (2008). Damage and degradation rates of extracellular DNA in marine sediments: Implications for the preservation of gene sequences. *Molecular Ecology*, *17*(17), 3939–3951. https://doi.org/10.1111/j.1365-294X.2008.03880.x

Cristescu, M. E. (2019). Can Environmental RNA Revolutionize Biodiversity Science? *Trends in Ecology and Evolution*, *34*(8), 694–697. https://doi.org/10.1016/j.tree.2019.05.003

Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 209–230. https://doi.org/10.1146/annurev-ecolsys-110617-062306

Deiner, K., & 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., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, *7*, 12544. https://doi.org/10.1101/020800

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

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., …

569 Gardner, P. P. (2018). Towards robust and repeatable sampling methods in eDNA-
570 based studies. *Molecular Ecology Resources*, *18*(5), 940–952.
571 https://doi.org/10.1111/1755-0998.12907

572 Dowle, E., Pochon, X., Keeley, N., & Wood, S. A. (2015). Assessing the effects of salmon
573 farming seabed enrichment using bacterial community diversity and high-throughput
574 sequencing. *FEMS Microbiology Ecology*, *91*(8), fiv089.
575 https://doi.org/10.1093/femsec/fiv089

576 Guardiola, M., Wangensteen, O. S., Taberlet, P., Coissac, E., Uriz, M. J., & Turon, X. (2016).
577 Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment
578 DNA and RNA. *PeerJ*, *4*, e2807. https://doi.org/10.7717/peerj.2807

579 Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., … Winfield, I. J.
580 (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term
581 data from established survey methods. *Molecular Ecology*, *25*(13), 3101–3119.
582 https://doi.org/10.1111/mec.13660

583 Harper, K. J., Anucha, N. P., Turnbull, J. F., Bean, C. W., & Leaver, M. J. (2018). Searching
584 for a signal: Environmental DNA (eDNA) for the detection of invasive signal crayfish,
585 *Pacifastacus leniusculus* (Dana, 1852). *Management of Biological Invasions*, *9*(2), 137–
586 148.

587 Hartig, F. (2021). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed)
588 Regression Models. Retrieved from http://florianhartig.github.io/DHARMa/

589 Hasnain, S. S., Escobar, M. D., & Shuter, B. J. (2018). Estimating thermal response metrics
590 for North American freshwater fish using Bayesian phylogenetic regression. *Canadian*
591 *Journal of Fisheries and Aquatic Sciences*, *75*(11), 1878–1885.
592 https://doi.org/10.1139/cjfas-2017-0278

593 Jane, S. F., Wilcox, T. M., Mckelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., …
594 Whiteley, A. R. (2015). Distance, flow and PCR inhibition: eDNA dynamics in two
595 headwater streams. *Molecular Ecology Resources*, *15*(1), 216–227.
596 https://doi.org/10.1111/1755-0998.12285

597 Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-unseen”
598 detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2),
599 150–157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

600 Jerde, C. L., Olds, B. P., Shogren, A. J., Andruszkiewicz, E. A., Mahon, A. R., Bolster, D., &
601 Tank, J. L. (2016). Influence of Stream Bottom Substrate on Retention and Transport of
602 Vertebrate Environmental DNA. *Environmental Science and Technology*, *50*(16), 8770–
603 8779. https://doi.org/10.1021/acs.est.6b01761

604 Jeunen, G. J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., …
605 Gemmell, N. J. (2019). Environmental DNA (eDNA) metabarcoding reveals strong
606 discrimination among diverse marine habitats connected by water movement.
607 *Molecular Ecology Resources*, *19*(2), 426–438. https://doi.org/10.1111/1755-
608 0998.12982

609 Laroche, O., Wood, S. A., Tremblay, L. A., Ellis, J. I., Lear, G., & Pochon, X. (2018). A cross-
610 taxa study using environmental DNA / RNA metabarcoding to measure biological
611 impacts of offshore oil and gas drilling and production operations. *Marine Pollution*
612 *Bulletin*, *127*, 97–107. https://doi.org/10.1016/j.marpolbul.2017.11.042

Laroche, O., Wood, S. A., Tremblay, L. A., Lear, G., Ellis, J. I., & Pochon, X. (2017). Metabarcoding monitoring analysis: the pros and cons of using co-extracted environmental DNA and RNA data to assess offshore oil production impacts on benthic communities. *PeerJ*, *5*, e3347. https://doi.org/10.7717/peerj.3347

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

Li, Y., & Breaker, R. R. (1999). Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2′-Hydroxyl Group. *Journal of the American Chemical Society*, *121*(4), 5364–5372.

Littlefair, J. E., Hrenchuk, L. E., Blanchfield, P. J., Rennie, M. D., & Cristescu, M. E. (2021). Thermal stratification and fish thermal preference explain vertical eDNA distributions in lakes. *Molecular Ecology*, *30*(13), 3083–3096. https://doi.org/10.1111/mec.15623

[dataset] Littlefair, J. E., Rennie, M. D., & Cristescu, M. E.; (2021); Environmental nucleic acids: a field-based comparison for monitoring freshwater habitats using eDNA and eRNA; Dryad; Version 1; DOI upon article acceptance

Marshall, N. T., Vanderploeg, H. A., & Chaganti, S. R. (2021). Environmental (e)RNA advances the reliability of eDNA by predicting its age. *Scientific Reports*, *11*, 2769. https://doi.org/10.1038/s41598-021-82205-4

Mee, J. A., & Taylor, E. B. (2012). The cybrid invasion: widespread postglacial dispersal by *Phoxinus* (Pisces: Cyprinidae) cytoplasmic hybrids. *Canadian Journal of Zoology*, *90*(5), 577–584. https://doi.org/10.1139/z2012-023

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., … Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*(7), 150088. https://doi.org/10.1098/rsos.150088

Miyata, K., Inoue, Y., Amano, Y., Nishioka, T., Yamane, M., Kawaguchi, T., … Honda, H. (2021). Fish environmental RNA enables precise ecological surveys with high positive predictivity. *Ecological Indicators*, *128*, 107796. https://doi.org/10.1016/j.ecolind.2021.107796

Nicholls, S. M., Quick, J. C., Tang, S., & Loman, N. J. (2019). Ultra-deep, long-read nanopore sequencing of mock microbial community standards. *GigaScience*, *8*(5), giz043. https://doi.org/10.1093/gigascience/giz043

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Helene, W. (2019). vegan: Community Ecology Package. R package version 2.5-6. Retrieved from https://cran.r-project.org/package=vegan

Orsi, W., Biddle, J. F., & Edgcomb, V. (2013). Deep Sequencing of Subseafloor Eukaryotic rRNA Reveals Active Fungi across Marine Subsurface Provinces. *PloS One*, *8*(2), e56335. https://doi.org/10.1371/journal.pone.0056335

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T., & Wilding, T. A. (2014). Environmental monitoring through protist next-generation sequencing metabarcoding: Assessing the impact of fish farming on benthic foraminifera communities. *Molecular*

Ecology Resources, *14*(6), 1129–1140. https://doi.org/10.1111/1755-0998.12261

Pawlowski, J., Esling, P., Lejzerowicz, F., Cordier, T., Visco, J. A., Martins, C. I. M., … Cedhagen, T. (2016). Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. *Aquaculture Environment Interactions*, *8*, 371–386. https://doi.org/10.3354/AEI00182

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., … Kahlert, M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, *637–638*, 1295–1310. https://doi.org/10.1016/j.scitotenv.2018.05.002

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

Pochon, X., Zaiko, A., Fletcher, L. M., Laroche, O., & Wood, S. A. (2017). Wanted dead or alive? Using metabarcoding of environmental DNA and RNA to distinguish living assemblages for biosecurity applications. *PLoS ONE*, *12*(11), e0187636. https://doi.org/10.1371/journal.pone.0187636

Rennie, M. D., Kennedy, P. J., Mills, K. H., Podemski, C. L., Rodgers, C. M. C., Charles, C., … Paterson, M. J. (2019). Impacts of freshwater aquaculture on fish communities: A whole-ecosystem experimental approach. *Freshwater Biology*, *64*(5), 870–885. https://doi.org/10.1111/fwb.13269

Sansom, B. J., & Sassoubre, L. M. (2017). Environmental DNA (eDNA) Shedding and Decay Rates to Model Freshwater Mussel eDNA Transport in a River. *Environmental Science and Technology*, *51*(24), 14244–14253. https://doi.org/10.1021/acs.est.7b05199

Seymour, M., Durance, I., Cosby, B. J., Deiner, K., Ormerod, S. J., Colbourne, J. K., … Biology, E. (2018). Acidity promotes degradation of multi-species eDNA in lotic mesocosms. *Communications Biology*, *1*, 4. https://doi.org/10.1038/s42003-017-0005-3

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., & Bolster, D. (2017). Controls on eDNA movement in streams: Transport, Retention, and Resuspension. *Scientific Reports*, *7*, 5065. https://doi.org/10.1038/s41598-017-05223-1

Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M., & Harvey, E. S. (2018). Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology*, *33*(1), 196–205. https://doi.org/10.1111/cobi.13183

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

Tan, S. C., & Yiap, B. C. (2009). DNA, RNA, and Protein Extraction: The Past and The Present. *Journal of Biomedicine and Biotechnology*, *2009*, 574398. https://doi.org/10.1155/2009/574398

Tavares, L., Alves, P. M., Ferreira, R. B., & Santos, C. N. (2011). Comparison of different methods for DNA-free RNA isolation from SK-N-MC neuroblastoma. *BMC Research Notes*, *4*, 3. https://doi.org/10.1186/1756-0500-4-3

699 Tillotson, M. D., Kelly, R. P., Duda, J. J., Hoy, M., Kralj, J., & Quinn, T. P. (2018).
700 Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at
701 fine spatial and temporal scales. *Biological Conservation*, *220*(July 2017), 1–11.
702 https://doi.org/10.1016/j.biocon.2018.01.030

703 Torti, A., Lever, M. A., & Jørgensen, B. B. (2015). Origin, dynamics, and implications of
704 extracellular DNA pools in marine sediments. *Marine Genomics*, *24*(3), 185–196.
705 https://doi.org/10.1016/j.margen.2015.08.007

706 Truelove, N. K., Andruszkiewicz, E. A., & Block, B. A. (2019). A rapid environmental DNA
707 method for detecting white sharks in the open ocean. *Methods in Ecology and*
708 *Evolution*, *10*(8), 1128–1135. https://doi.org/10.1111/2041-210X.13201

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

712 Visco, J. A., Apothéloz-Perret-Gentil, L., Cordonier, A., Esling, P., Pillet, L., & Pawlowski, J.
713 (2015). Environmental Monitoring: Inferring the Diatom Index from Next-Generation
714 Sequencing Data. *Environmental Science and Technology*, *49*(13), 7597–7605.
715 https://doi.org/10.1021/es506158m

716 Wang, Y., Naumann, U., Wright, S. T., & Warton, D. I. (2012). Mvabund- an R package for
717 model-based analysis of multivariate abundance data. *Methods in Ecology and*
718 *Evolution*, *3*(3), 471–474. https://doi.org/10.1111/j.2041-210X.2012.00190.x

719 Wood, S. A., Biessy, L., Latchford, J. L., Zaiko, A., Ammon, U. Von, Audrezet, F., … Pochon,
720 X. (2020). Release and degradation of environmental DNA and RNA in a marine
721 system. *Science of the Total Environment*, *704*, 135314.
722 https://doi.org/10.1016/j.scitotenv.2019.135314

723 Yates, M. C., Derry, A. M., & Cristescu, M. E. (2021). Environmental RNA: A Revolution in
724 Ecological Resolution? *Trends in Ecology & Evolution*, *36*(7), 601–609.
725 https://doi.org/10.1016/j.tree.2021.03.001

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

730

731 Data Accessibility statement:

732 Raw fastq files, sample x ASV tables, and the sequence composition of the ASVs are openly
733 available at Dryad (xxx). Scripts to process bioinformatic data are available from
734 https://github.com/CristescuLab/YAAP.

735 Author contributions:

736 JEL and MEC designed the research, all authors contributed to funding the research, JEL
737 performed the molecular work and analysis, MDR contributed field data, JEL wrote the first
738 draft of the manuscript and all authors contributed the writing and editing of the manuscript.

739 Table 1: Sequences present at each step of bioinformatic filtering.

740

741

742

Bioinformatics step	**Mean ± 95% confidence error (per sample)**	
	DNA	**RNA**
Initial sequence number	83,972 ± 11,995	78,751 ± 12,075
Initial sequence quality score	32.5 ± 0.61	32.0 ± 0.61
Adapter trim (Cutadapt)	74,493 ± 11134	68,107 ± 11,326
Merge reads (Pear)	74,450 ± 11131	67,169 ± 11,397
Trim of remaining reads (Cutadapt)	74,444 ± 11130	68,018 ± 11,318
Length filter	72,636 ± 11112	50,735 ± 9,932
Denoising pipeline	**Mean ± 95% confidence error (all samples)**	
Total sequences after length filtering	7,118,369	4,971,983
Number of non-chimeric ASVs	107	115
Number of reads mapped onto ASVs	7,012,122 of 7,118,369 = 98.5%	4,890,593 of 4,971,983 = 98.4%

Figure 1: Species accumulation curves for each lake habitat in the study, based on incidence data. The dataframe was filtered to include only fish species present in the northwest Ontario region (i.e. removing non-target ASVs). Separate curves are drawn for ASVs detected with DNA and RNA molecules in the two sampling seasons (August and October), created with random sampling of sites in function specaccum in "vegan". The grey dashed line indicates the expected number of fish species in that lake according to conventional fishing methods.

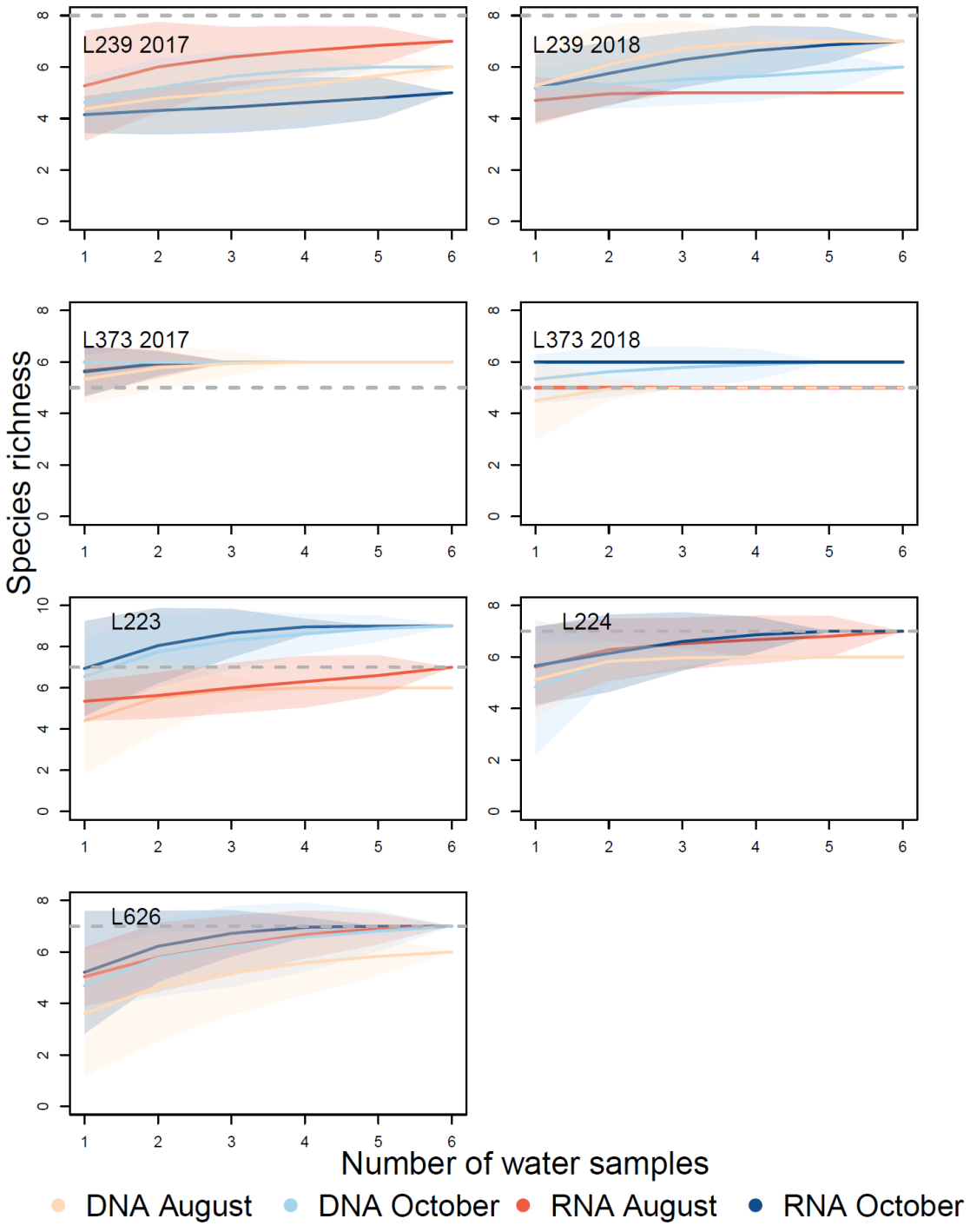


Figure 2: A) True positive rate and B) false discovery rate as measured by eDNA and eRNA. eRNA had a significantly higher true positive rate than eDNA. True positive rate is defined as the proportion of true positives out of true positives and false negatives combined. There was no significant difference in false discovery rate between eDNA and eRNA. False discovery rate is defined as the proportion of false positives out of all detections (i.e. true positives and false positives together). Error bars are standard errors of the mean.

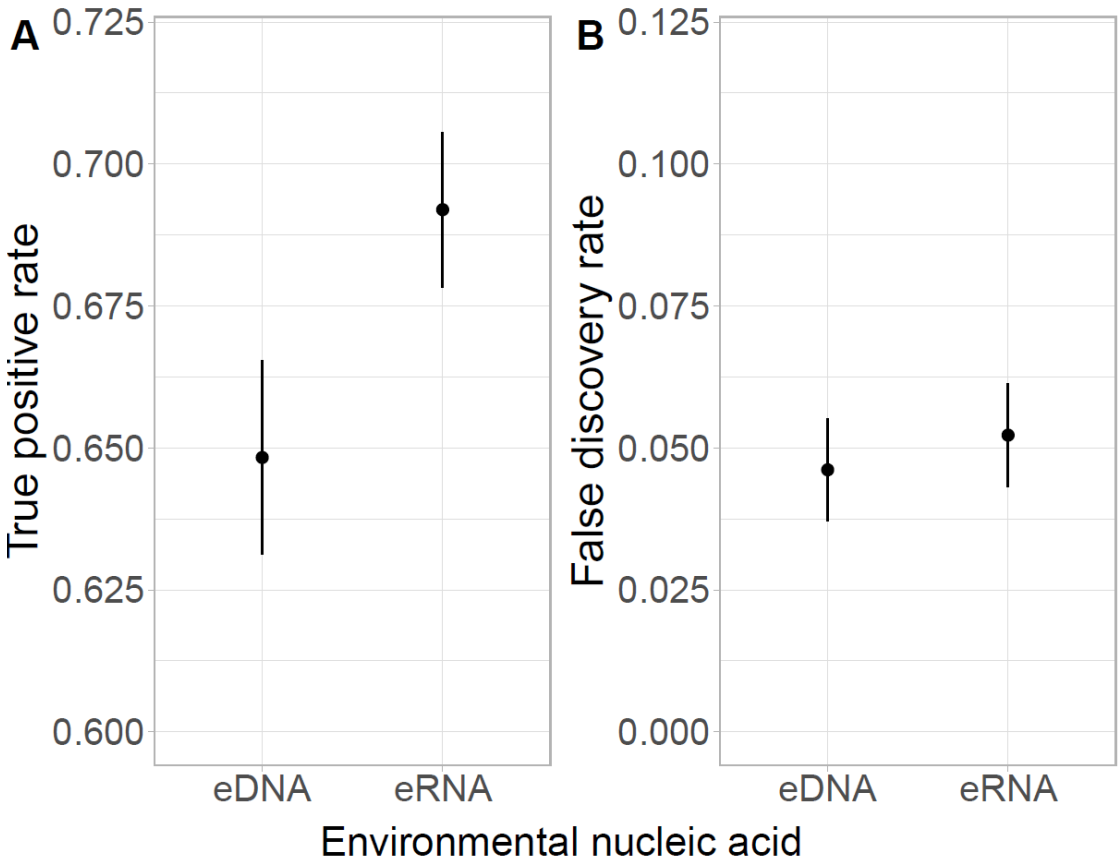
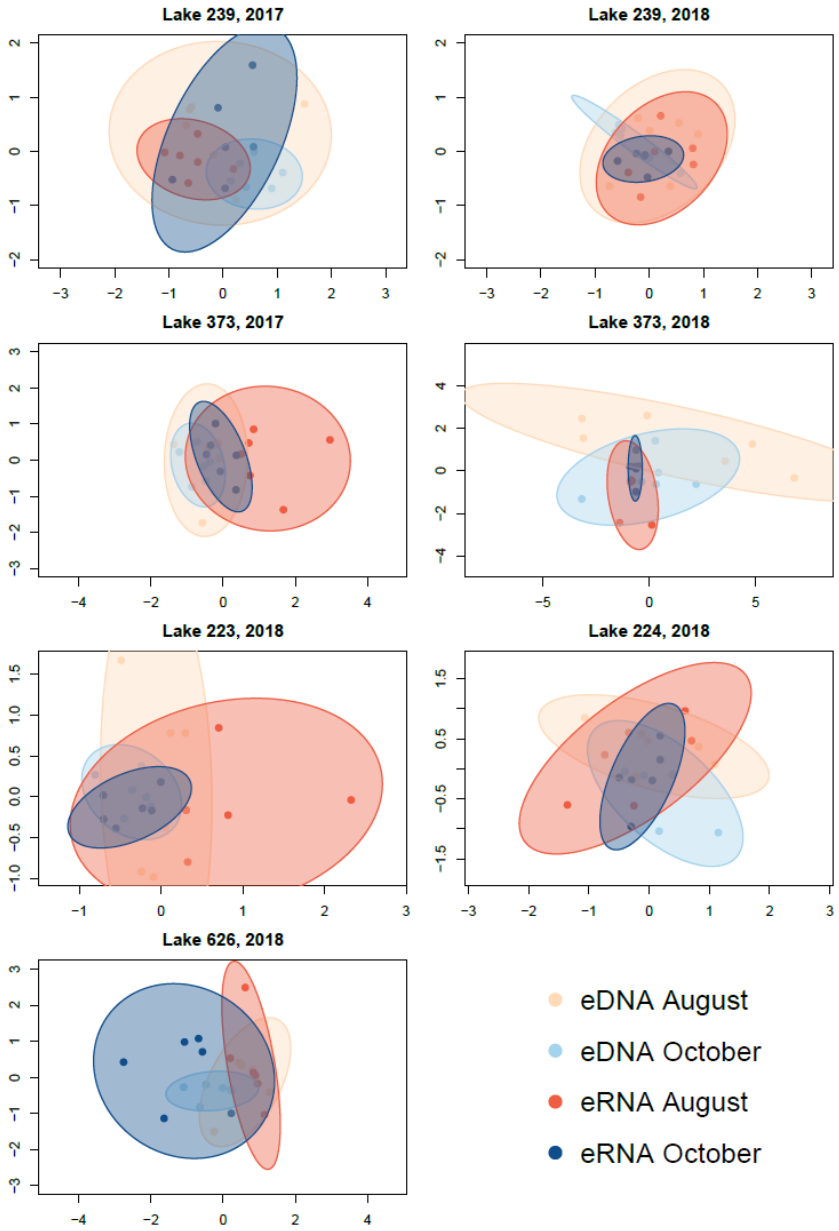
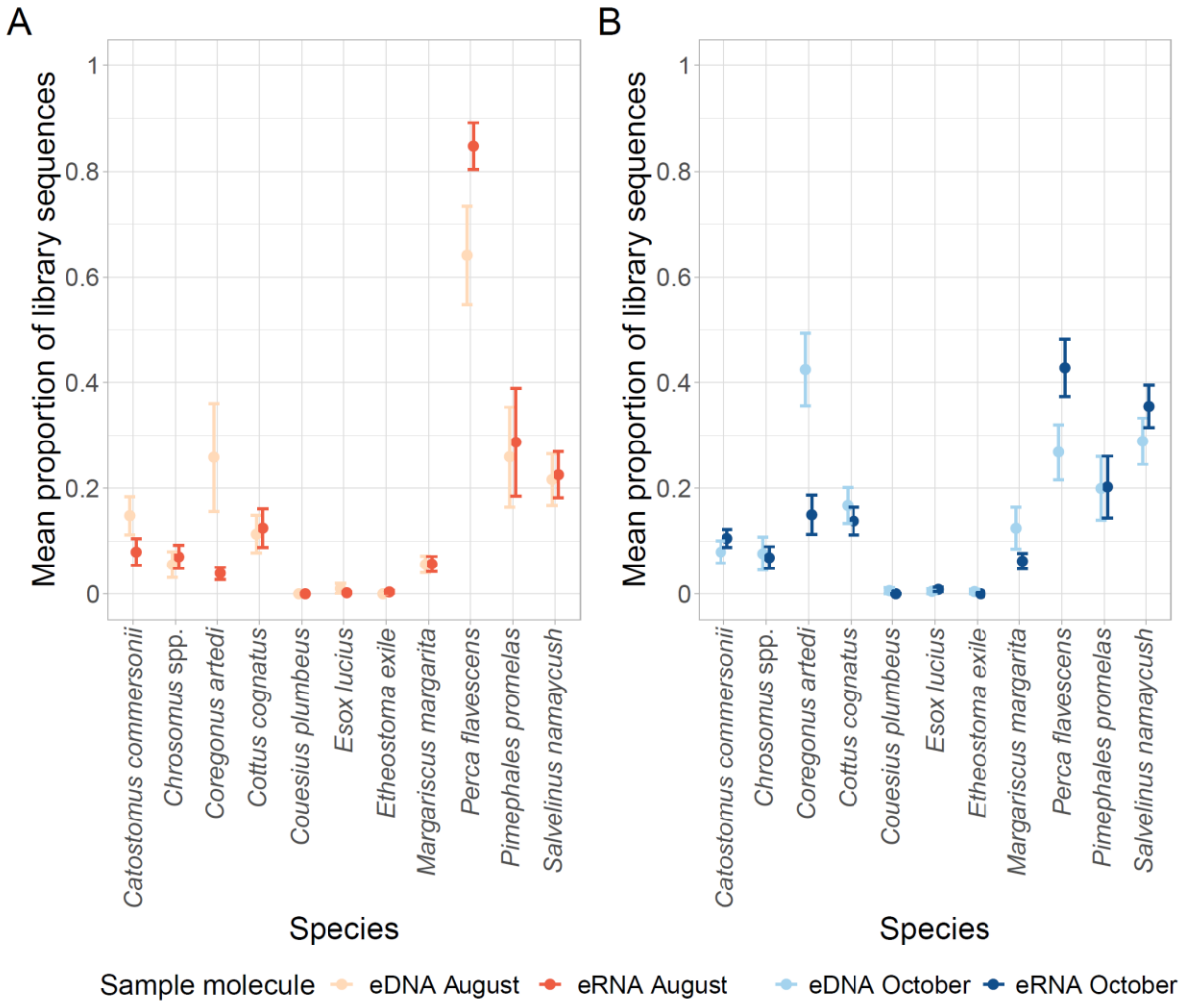


Figure 3: nMDS plots for each lake showing community dissimilarities for both eDNA and eRNA in August and October. New nMDS are run for each lake; thus, scaling varies for each image. Ellipses are 95% confidence intervals coloured according to environmental nucleic acid and season.



807 Figure 4: Proportional composition of ASVs per filter for each fish species for A) August and
808 B) October. Where there is more than one ASV per species, these are grouped together
809 under the species name. Error bars are standard errors of the mean.

810



811

812

813

814