# Executing multi-taxa eDNA ecological assessment via traditional metrics and interactive networks

Mathew Seymour [a,*], Francois K Edwards [b], Bernard J Cosby [c], Martyn G Kelly [d], Mark de Bruyn [e], Gary R Carvalho [a], Simon Creer [a]

[a] Molecular Ecology and Fisheries Genetics Laboratory, School of Biological Sciences, Bangor University, Bangor, Gwynedd LL57 2UW, UK
[b] Centre for Ecology & Hydrology, Wallingford OX10 8BB, UK
[c] NERC Centre for Ecology & Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
[d] Bowburn Consultancy, 11 Monteigne Drive, Bowburn, Durham DH6 5QB, UK
[e] School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia

## HIGHLIGHTS

- Environmental DNA biological quality elements enable effective environmental assessment using high throughput sequencing
- Non-traditional 18S biological indicators provide a finer scale assessment of environmental conditions
- Network based analyses enable quantification of changes in biotic and abiotic associations across multiple taxonomic groups
- Transitivity network properties enable direct assessment with currently existing environmental policy metrics

## GRAPHICAL ABSTRACT
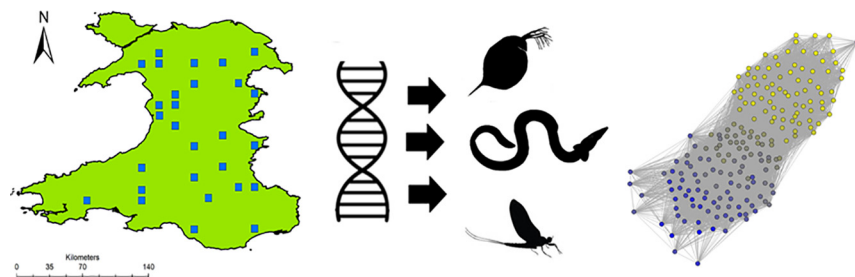


## ARTICLE INFO

## ABSTRACT

Current approaches to ecological assessment are limited by the traditional morpho-taxonomic methods presently employed and the inability to meet increasing demands for rapid assessments. Advancements in high throughput sequencing now enable rapid high-resolution ecological assessment using environmental DNA (eDNA). Here we test the ability of using eDNA-based ecological assessment methods against traditional assessment of two key indicator groups (diatoms and macroinvertebrates) and show how eDNA across multiple gene regions (COI, rbcL, 12S and 18S) can be used to infer interactive networks that link to ecological assessment criteria. We compared results between taxonomic and eDNA based assessments and found significant positive associations between macroinvertebrate ($p < 0.001$ R$^2$ = 0.645) and diatom ($p = 0.015$, R$^2$ = 0.222) assessment metrics. We further assessed the ability of eDNA based assessment to identify environmentally sensitive genera and found an order of magnitude greater potential for 18S, versus COI or rbcL, to determine environmental filtering of ecologically assessed communities. Lastly, we compared the ability of traditional metrics against co-occurrence network properties of our combined 18S, COI and rbcL indicator genera to infer habitat quality measures currently used by managers. We found that transitivity (network connectivity), linkage density and cohesion were significantly associated with habitat modification scores (HMS), whereas network properties were inconsistent with linking to the habitat quality score (HQS) metric. The incorporation of multi-marker eDNA network assessment opens up a means for finer scale ecological assessment, currently limited using traditional methods. While utilization of eDNA-based assessment is recommended, direct comparisons with traditional

* Corresponding author.
E-mail address: mat.seymour@gmail.com (M. Seymour).

approaches are difficult as the methods are intrinsically different and should be treated as such with regards to future research. Overall, our findings show that eDNA can be used for effective ecological assessment while offering a wider range of scope and application compared to traditional assessment methods.

## 1. Introduction

A major challenge for the 21st century is ensuring natural resources are well managed for current and future generations (Lampert, 2019). Increased awareness of the negative impacts of human interactions, paired with continued technological development and budgetary constraints, calls for a practical reassessment of how we implement environmental regulation and the strengthening of links between assessment methods and underlying ecological principles. Freshwater ecosystem management is immensely important as access to freshwater via lakes and rivers is essential for human societies. Freshwater ecosystems are highly sensitive and are regularly exposed to multiple stressors including toxic chemicals, pharmaceuticals, and thermal profiles, which compromise their stability and ecosystem function. Instability of natural waterways imposes a negative impact on human livelihood, particularly with regards to growth, development and sustainment of our populations and communities. Whereas direct chemical assessment would intuitively seem the most logical means to evaluate natural systems for contamination, it often lacks temporal and spatial reliability resulting in false positive or negative associations (Morse et al., 2007). Additionally, with the ever increasing number of manufactured compounds, the ability to detect all possible point effects and the innumerable interactions of potential environmental stressors is impractical (Gavrilescu et al., 2015). Ecological assessment, whereby observations of living populations or communities within and among sampling locations is used to assess natural resource health, is a well-established practice and serves as the basis for freshwater regulation such as that required by the Water Framework Directive (WFD) (Morse et al., 2007).

The WFD is a key piece of legislation within the European Union (EU) with the primary objective of ensuring sustainable aquatic resources for future generations. Implementation of the WFD requires a system for evaluating the health of aquatic ecosystems, as a means of identifying those in need of remediation, and to ensure no deterioration of ecosystems currently in a healthy state (referred to as "good ecological status") (Carvalho et al., 2019). The WFD requires characterization of biological communities, along with physiochemical and hydromorphological conditions. The ecological status of a given freshwater body is determined by the assessment of biological quality elements (BQEs); including fish, benthic macroinvertebrates, phytobenthos, macrophytes and phytoplankton, each requiring unique sampling, analysis and computational approaches (e.g. metrics) (Simboura et al., 2005). Benthic macroinvertebrate metrics, such as the Whalley, Hawkes, Paisley & Trigg (WHPT) metric (Water Framework Directive, 2014b), function by assigning taxonomically identifiable groups to scores based on their pollution tolerance, which are then used to generate ecological assessment metrics, including Average Score per Taxon (ASPT) (Mandaville, 2002), from which ecological status can be derived. Similarly, the Trophic Diatom Index (TDI) assigns scores based on pollution tolerance to diatom species (Kelly, 1998). Furthermore, current biomonitoring metrics rely heavily on the assumption of environmental sorting of communities, which may induce false positive or negative associations of the sampled individuals if other community factors are not considered, including species interactions and dispersal dynamics (Cadotte and Tucker, 2017). While increased replication would allow for greater accuracy under traditional protocols, current budgetary and manpower constraints hinder effective implementation.

Traditional biomonitoring protocols also rely heavily on identification of organisms by skilled analysts and can be time and resource intensive if a centralized identification database is not used to cross reference variants (Baird and Hajibabaei, 2012). Alternatively, utilizing a molecular based sampling and taxonomic assignment protocol that relies on a centralized taxon database and computer assignment may allow for a standardized biomonitoring protocol as well as furthering the scope and depth of the indicator groups used to associate pollution or environmental perturbation, thereby allowing finer scale assessment of any environmental changes that have occurred at target sites (Pawlowski et al., 2018).

Environmental DNA (eDNA) is DNA isolated from environmental samples and has been shown to be a reliable means for assessing biodiversity in freshwater environments (Bohmann et al., 2014; Seymour, 2019). Utilizing eDNA for community biomonitoring is not yet fully recognized, however eDNA based methods are increasingly being used across government and industry sections to assess individual species populations that are of conservation concern (Biggs et al., 2015), or pose a threat as invasive species (Jerde et al., 2013). Environmental DNA assessment methods facilitate rapid sampling and can be analyzed faster than traditional methods, thereby providing greater spatial and replication scope when sampling a given environment. The eDNA-based assessment approach also facilitates a standardized sampling and assessment protocol across broad taxonomic space, providing several orders of magnitude more diversity data generated per sample than traditional assessment methods. With regards to traditional community biomonitoring, eDNA-based assessment still needs to be validated as a viable surrogate for current biomonitoring protocols, but offers several potential advantages with increased sensitivity, spatial and temporal scope, and reduced manpower requirements (Cordier et al., 2017; Pawlowski et al., 2018). Currently, eDNA based biomonitoring studies have shown good agreement between traditional and eDNA based methods (Gibson et al., 2015), but many are limited in their spatial scale, having been restricted to mesocosm experiments often with narrow taxonomic breadth, and do not link to an existing ecological assessment strategy (though see recent efforts regarding diatom eDNA biomonitoring; Vasselon et al., 2018).

In contrast to current biomonitoring data, environmental DNA-derived data also has the potential to include a much wider range of taxa and indicator groups that are not currently included due to limitations of traditional taxonomic identification. This may allow for a finer scale assessment, particularly in assessing differences among neighboring sites or in evaluating moderate changes in environmental conditions (Cordier et al., 2017). While there is a need to ensure continuity between traditional and eDNA based assessment methods, there should also be a forward thinking approach to utilizing the broader eDNA data to derive comparable biometric data that encompass not only the concept of environmental filtering on individual indicator taxa, but also the interactive qualities that are also linked to the stability and ecological complexity of habitats (Cadotte and Tucker, 2017). Co-occurrence networks are one way to assess taxa interactions across deep diversity pools, and have been linked to several aspects of metacommunity dynamics that are directly relatable to biomonitoring principles (Araújo et al., 2011). From co-occurrence networks, several network properties can be used to summarize the interactive relationship between abiotic (effects of the environment on the individual indicators) and the biotic (strength of the relationship between sets of indicators). Linkage density is a surrogate measure of community complexity, whereas transitivity provides a measure of network

connectivity. Modularity indicates network redundancy, and cohesion indicates the sensitivity of the community to changes in community structure (Karimi et al., 2017).

This study assesses the reliability of applying eDNA based ecological assessment metrics using matched traditional and eDNA based sampling from a national biomonitoring program. We compared biomonitoring assessment approaches for benthic macroinvertebrates and diatoms using both traditional and eDNA-derived data to determine whether eDNA can be used as a surrogate for traditional methods. We further assessed the ability of using eDNA genera groups as indicators by comparing environmental association between the 18S nuclear small subunit (nSSU) universal molecular marker and molecular markers that targets traditional biomonitoring groups (rbcL, 12S and COI). The expectation being that traditional biomonitoring groups are the preferred monitoring biological units to determine inter-site variations based on environmental differences, since biomonitoring methods were originally designed for the traditional indicator groups. Finally, we assessed the ability of co-occurrence networks of environmentally sensitive genera to infer ecological status of rivers, with the expectation that network properties associated with connectivity among genera (i.e. network vertices) should link to differences in ecological status.

## 2. Methods

### 2.1. Sampling

The Glastir Monitoring and Evaluation Programme (GMEP) was a Wales-wide environmental assessment program that included assessment of freshwater biodiversity (Emmett et al., 2017). The GMEP assessments were based on structured, unbiased reporting of ongoing national trends in widespread habitat, soil and landscape types across the rural and peri-urban landscape of Wales, based on a survey of 300 squares (1 × 1 km) sampled over a 4-year period between 2012 and 2016. The survey squares used for sampling were chosen by random sampling within Institute of Terrestrial Ecology (ITE) land classes to provide a good representation of widespread broad habitats and wider countryside. Headwater streams are typically found in 50–60% of the randomly selected GMEP squares. In the 75 squares sampled in 2016, 35 included rivers for which eDNA samples were collected, and eDNA successfully analyzed for 31 of the collected samples for this study (Fig. 1). The streams sampled for eDNA were also analyzed for chemical,



**Fig. 1.** Map of Wales and the GMEP sampling sites (blue squares).

physical and biological metrics including diatoms, macrophytes, macroinvertebrates and habitat suitability. For each sampling event (i.e. site) we collected eDNA and traditional invertebrate kick-net samples on the same day. Each site was visited on a separate day with separate sampling kits to avoid contamination among sites. Kits consisted of single replicate sampling material prepared in sterile plastic bags and have been found to be highly effective in avoiding cross contamination in a previous 300 replicate experiment (Seymour et al., 2018). Water samples for eDNA analysis (1 L) were collected, in triplicate, from each stream and filtered through enclosed 0.22 μm Sterivex filter units (EMD Millipore Corporation, Billerica, USA) using a Geopump TM Series II peristaltic pump (Geotech, Denver, USA). Filters were immediately preserved in lysis buffer (buffer ATL; Qiagen, Venlo, The Netherlands) and stored at room temperature during transit to the laboratory for further processing. Invertebrate communities were sampled using a standardized 3-minute kick sampling protocol, with a kick/hand net (500 μm mesh gauge). Both bank margins and riffle habitats were sampled during this timed sampling period. Invertebrates were preserved in absolute ethanol (99.8%; VWR International, Lutterworth, UK) after collection. On return to the laboratory, invertebrates were sorted from other collected material and identified to the lowest practical level. Benthic diatom samples were collected from submerged surface structures; cobble when available, otherwise large pebbles and small boulders were used. At least five separate surface structures with obvious diatom film were sampled per site. Structures were placed in a tray with 50 mL of stream water and brushed vigorously with a stiff toothbrush. After each structure had been scrubbed in the same tray, the suspension was transferred to sample bottles for transport back to the lab for identification.

### 2.2. Extraction and sequencing

DNA from filters was extracted using a modified QIAGEN DNA blood and tissue extraction protocol (Spens et al., 2017). Extracts were then cleaned for impurities using QIAGEN Power Clean kit and frozen at −20° for subsequent analyses. Four sequence libraries were created using a twostep protocol (e.g. Bista et al., 2017) with unique matching index ends (manufactured by Integrated DNA Technologies, IDT) and the following metabarcode gene regions: 1) 18S primers TAReuk454FWD1 (5′- 463 CCAGCA(G/C)C(C/T) GCGGTAATTCC-3′) and TAReukREV3 (5′- 464 ACTTTCGTTCTTGAT (C/T)(A/G)A-3′) (Stoeck et al., 2010); 2) COI primers m1COIintF (5′-GGWACWGGWTGAACWGTWTAYCCYCC-3′) and jgHCO2198 (5′-TAIACYTCIGGRTGICCRAARAAYCA-3′) (Leray et al., 2013); 3) rbcL rbcL646F (5′-ATGCGTTGGAGAGARCGTTTC-3) and rbcL998R (5′-GATCACCTTCTAATTTACCWACAACTG-3) (Glover, 2019); 4) 12S MiFish-U forward (5′-GTCGGTAAAACTCGTGCCAGC-3) and reverse (5′-CATAGTGGGGTATCTAATCCCAGTTTG-3) (Miya et al., 2015). Each sample library was amplified using each primer set in triplicates which were then pooled and index labeled with assistance of a Gilson pipette max liquid handler. For the 18S, rbcL and 12S barcodes we used New England Biolab's Q5 mastermix. For the COI barcodes we utilized Thermo Scientific's Ampli-gold mastermix, due to the higher number of inosine in the COI barcodes, which would not amplify when using the Q5 master mix. For the 2nd round index PCR, Q5 master mix was used for all reactions. All PCRs were run in 25 μL reactions. For 18S barcodes, PCR reactions consisted of 1.5 μL DNA template, 12.5 μL mastermix, 1 μL (10 ng/ μL)of each primer pair and 9 μL PCR grade water, and amplified using 98 °C for 30 s followed by 15 cycles of 98 °C for 10 s, 50 °C for 30 s and 72 °C for 30 s followed by a final annealing step at 72 °C for 10 min. COI barcodes were generated using a reaction mix of 12.5 μL mastermix, 2 μL DNA template, 1 μL of each primer and 8 μL nuclease free water and amplified using an initial 95 °C for 5 min then 25 cycles of 95 °C for 30 s, 54 °C for 30 s and 72 °C for 60 s followed by a 72 °C final annealing for 10 min. RbcL barcode
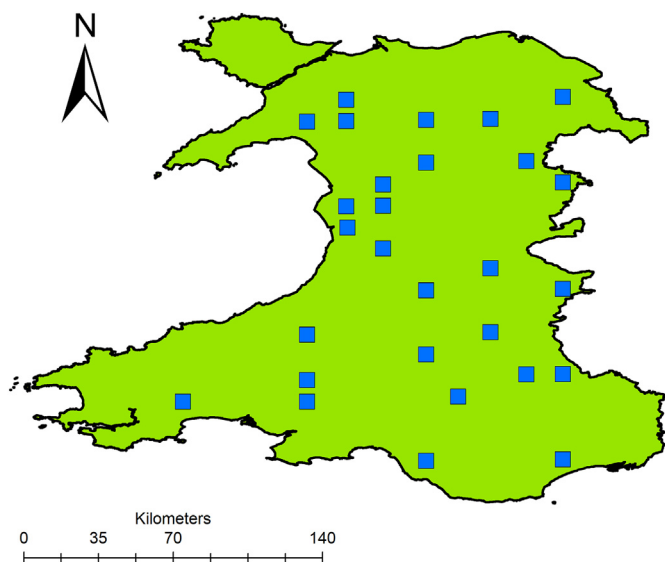
reactions were setup using 12.5 μL master mix, 1.5 μL DNA template and 9 μL nuclease free water and amplified using 98 °C at 30 s with 23 cycles of 98 °C for 5 s, 54 °C for 10 s and 72 °C for 20 s followed by 72 °C for 5 min. 12S barcode reactions were initiated with 12.5 μL mastermix, 2 μL DNA template, 1 μL of each primer and 9 μL nuclease free water and amplified using 98 °C for 30 s, followed by 15 cycles of 98 °C for 10 s, 54 °C for 30 s, 72 °C for 30 s with a final annealing step of 72 °C for 5 min. Libraries were then shipped to the University of Birmingham's Genomic sequencing facility for quality control and sequencing. For quality control amplicons were purified using High Prep PCR magnetic beads (Auto Q Biosciences) and quantitated using a 200 pro plate reader (TECAN) using qubit dsDNA HS solution (Invitrogen). Final library amplicons were mixed in equimolar quantities (at a final concentration of 12 pmol) using a Biomek FXp liquid handling robot (Beckman Coulter). Pool molarity was confirmed using a HS D1000 tapestation screentape (Agilent) prior to $2 \times 250$ bp paired end sequencing on an Illumina HiSeq platform to obtain ≥100,000 reads per sample.

### 2.3. Bioinformatics

Bioinformatics up to taxonomic assignments were performed by the University of Birmingham. Amplicon reads were demultiplexed by the sequencing facility. Quality based trimming of the fastq reads were performed using SolexaQA++ v.3.1.7.1 (Cox et al., 2010). Paired ends were merged using Flash (Magoč and Salzberg, 2011) under the default parameters. Primer sequences were removed by TagCleaner, whereby up to 3 mismatches per primer sequence was allowed (Schmieder et al., 2010; Miya et al., 2020). Only sequences with both forward and reverse primers were retained for further analyses. Amplicon sequence variants (ASVs) were obtained using Unoise3 using default parameters (Edgar, 2016). Diatom taxonomy was assigned to representative ASVs using the Diat.barcode data based at 97% similarity threshold (Rimet et al., 2019). COI and 12S taxonomy was assigned to representative ASVs using the non-redundant nucleotide database (NCBI) at 97% similarity threshold (Alberdi et al., 2018). 18S taxonomy was assigned to representative ASVs using the SILVA v128 database at 97% similarity (Quast et al., 2013). Taxonomic assignments, and their associate ASVs that returned incomplete taxonomy or unknown identifiers were excluded from further analyses. Amplicon sequence variants were rarified to the lowest replicate level to normalize the diversity of the samples. Mean read numbers for each amplicon sequence variant (ASV) were calculated across the rarified sequences before being matched to their taxonomic identifier. Amplicon sequence variants that did not occur in at least two of the three site replicates were discarded from subsequent analyses.

### 2.4. Statistics

All statistical analyses were performed using R version 3.6.1 (R Core Team, 2019). Whalley, Hawkes, Paisley & Trigg (WHPT) average score per taxa (ASPT) was based on WFD river invertebrate metric scores to the identified family groups from the traditional and eDNA based sampling (Water Framework Directive, 2014a). WHPT is an abundance-based method that assigns metric scores based on raw or log scale abundance groups. To adapt the traditional metric to eDNA we used proportional read numbers within each sample to assign metric scores to each group, such that 0.1, 1, 10, >10 translated to 1, 10, 100, >100 class groupings in the traditional metric protocol (Water Framework Directive, 2014a). Correlations were assessed between traditional and eDNA based ASPT numbers for each site using type II linear regressions using a major axis regression. Diatom biomonitoring scores were calculated, following existing protocols, as the sum of the percentages for all taxa (counts for traditional sampling and read numbers for eDNA) in each sample using the TDI4 taxon scores for rivers (Kelly et al., 2008). Correlations were assessed between traditional and eDNA based TDI values

for each site using type II linear regressions using a major axis regression.

### 2.5. Networks

Co-occurrence networks were constructed in a two-step process that utilized the unique genera found across all four markers (Table S1). First, Genera sensitivity to environmental conditions across the sites was assessed against the first principle component axis of the PCA characterizing habitat variation across the sites (listed below). We used logistic regression and maximum likelihood to determine whether the presence or absence of each genera was significant ($p < 0.05$) against the response variable (first PCA axis characterizing environmental variation) across the sites. To account for unequal weighting due to study design sampling and unequal groupings of environmental variation across sites, an environmental weighting was used based on the density of the environmental values across the sites (King and Zeng, 2001). Additionally, each genus had to occur in at least 3 sites to be statistically tested. Second, genera co-occurrence (across all sites) networks were constructed from the genera who significantly ($p < 0.05$) associated with the PCA derived environmental gradient, with links between genera (edges) determined using the proportion of shared sites between each genera (vertices) as a threshold for linking (edges) genera. Due to the general nature of connectivity networks to be sensitive to threshold values, we took a holistic approach and calculated all networks for thresholds between 1% and 80% (0.5% increments; 159 resulting networks) co-occurrence, whereby connections (edges) between genera (vertices) were connected (non-directional) if they co-occurred at or above the threshold value (proportion of shared sites).

For each site we isolated the site-specific subnetwork ($N = 31$) for each of the co-occurrence networks created ($N = 159$). Subnetwork properties; including linkage density, modularity, transitivity and cohesion were calculated to compare changes in networks structure across sites for each threshold via the igraph package in R (Csardi and Nepusz, 2006). We assessed the relationship between each biomonitoring metric and network property vs the habitat quality score (HQS) and habitat modification score (HMS) using linear regression across all network co-occurrence threshold values, to assess overall relationships in subnetwork properties to water framework environmental quality measures.

## 3. Results

### 3.1. Environmental variation

Across all sites ($N = 31$) mean pH was 6.30 (SD = 0.81), conductivity was 106.5 μS/cm (SD = 116.85), alkalinity was 31.26 mg/L (SD = 51.36), width was 1.11 m (SD = 0.714), depth was 9.30 cm (SD = 4.88), boulder coverage was 36.55% (SD = 26.02), gravel coverage was 45.52% (SD = 21.27), sand coverage was 9% (SD = 7.17), silt coverage was 8.71% (SD = 20.76), moss coverage was 5.74% (SD = 8.75), algae coverage score was 1.71 (SD = 3.55), plant coverage score was 7.90 (SD = 18.96), altitude was 238.66 m (SD = 148.85) and slope was 84.27% (SD = 71.79). To avoid bias in the PCA analyses, environmental variables were assessed for significant correlation using Spearman tests, where variables that were highly correlated ($p < 0.01$) with several other variables were used as surrogate variables for their co-varying counterparts. As such, boulder coverage was surrogate for conductivity ($p < 0.001$, $R^2 = 0.613$), alkalinity ($p = 0.001$, $R^2 = 0.587$), width ($p < 0.001$, $R^2 = 0.599$), gravel ($p < 0.001$, $R^2 = 0.598$) and silt ($p < 0.001$, $R^2 = 0.645$). Depth was used as surrogate for moss ($p = 0.008$, $R^2 = 0.469$). This resulted in pH, depth, boulder, sand, algae coverage, altitude and slope being used to create the PCA based environmental gradient axes which in-turn explained 39.1% of the variation across the sites in the first PC axes.

## 3.2. Biodiversity

Average cleaned and rarified read numbers used for downstream analyses were 11,460 (SD = 6445) per site for COI, 46,981 (SD = 22,408) for rbcL, 101,036 (SD = 27,540) for 18S and 266 (SD = 680) for 12S. Average individuals used for kick-net taxonomy per site was 1179 (SD = 1388).

Across all sites we found 1157 unique universal (18S) genera, 219 unique macroinvertebrate (COI) genera, 138 unique diatom (rbcL) genera and 9 unique fish (12S) genera. We found an average of 66.12 macroinvertebrate genera across all sites (SD = 20.33) via COI eDNA, and 22.84 genera (SD = 8.57) with traditional kick netting. We found an average of 47.93 diatom (rbcL) genera (SD = 17.95) per site using eDNA and an average of 12.5 genera (SD = 4.67) per site using traditional sampling. The universal 18S marker found 412.23 genera (SD = 94.73) on average per site using eDNA. Fish (12S) diversity was on average 2 genera (SD = 1.43) across all sites using eDNA.
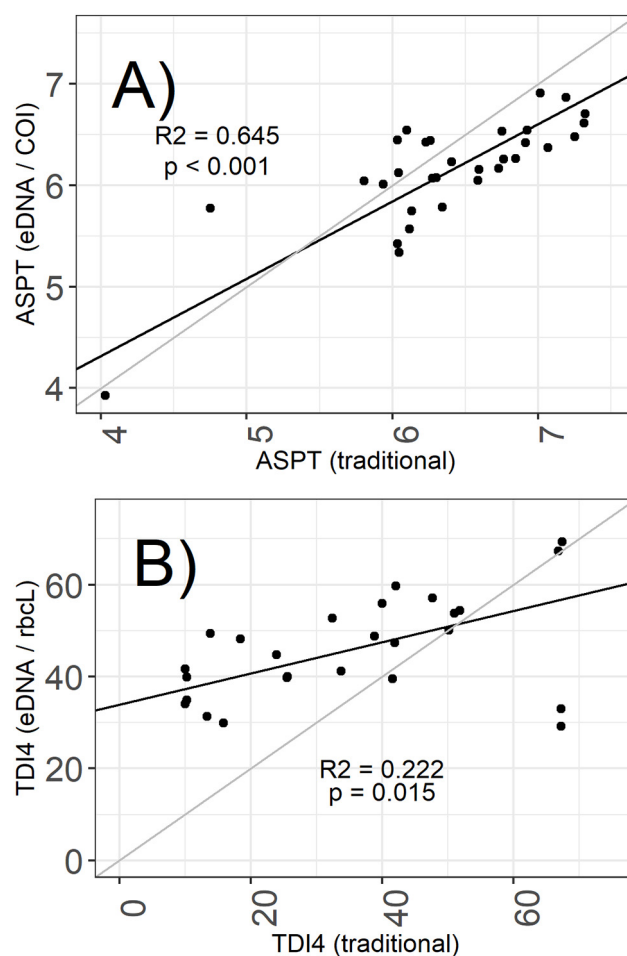
## 3.3. Bioscores

Across all sites, ASPT derived from macroinvertebrate eDNA was on average 6.140 (SD = 0.564) and 6.391 (SD = 0.702) using traditional surveying methods. Mean overlap in eDNA versus traditional biotic macroinvertebrate indicators was 47.83% (SD = 10.58) across all sites and 15.57% (SD = 5.97) for macroinvertebrate genera. Environmental DNA accounted for 28.46% (SD = 14.71) unique biotic indicators not detected by traditional methods and 64.01% (SD = 17.27) genera not detected by traditional methods. Conversely, 23.71% (SD = 17.17) of the macroinvertebrate biotic indicators were only detected with traditional methods and 20.44% (SD = 17.30) of macroinvertebrate genera. Diatom TDI4 was 44.671 (SD = 11.976) for eDNA derived and 35.290 (SD = 19.596) for traditional surveying. Mean overlap in eDNA versus traditional diatom indicators was 10.97% (SD = 6.85) and 33.44% (SD = 11.95) for genera. Environmental DNA accounted for 63.15% (SD = 16.36) unique diatom indicators and 49.33% (SD = 16.90) unique diatom genera. Traditional sampling accounted for 25.87% (SD = 17.01) unique diatom indicators and 17.23% (SD = 15.53) unique diatom genera.

ASPT significantly correlated (positive) between eDNA and traditional survey methods ($p < 0.001$, $R^2 = 0.645$). TDI4 was also significantly correlated (positive) between eDNA and traditional biomonitoring methods ($p = 0.015$, $R^2 = 0.222$) (Fig. 2).

## 3.4. Indicator genera assessment and co-occurrence networks

Indicator genera were identified from 18S, COI and rbcL assigned genera; however, no 12S assigned taxa were significantly associated with environmental variation among sites, and were excluded from further analyses. We found 214 indicator genera for 18S, 26 genera for COI and 12 for rbcL (Fig. 3). Main networks derived from the indicator genera varied in linkage across the threshold values ranging from 6 vertices at 80% co-occurrence to 249 at 0.1% (mean = 197, SD = 77). Network degree (the average connections per network vertex) ranged from 1.7 at 80% co-occurrence to 221.6 at 0.1% (mean 79.2, SD = 70.6). Across all networks, vertices (i.e. genera) with similar environmental sensitivity (Fig. 3) consistently clustered with genera of similar environmental sensitivity (Fig. 4).

Subnetwork properties (i.e. sites) varied across threshold values for modularity, transitivity, linkage density and cohesion (Fig. 5). Modularity was not significantly ($p > 0.05$) associated with HMS or HQA for any of the co-occurrence thresholds and remained close to 0 for most subnetworks (mean across threshold = 0.139, SD = 0.066), indicating random network associations. We found significant ($p < 0.05$) positive associations between transitivity and HMS for all levels between 17 and 50% as well as thresholds between 66 and 80%. Transitivity was not consistently associated with HQA but was significantly negatively



**Fig. 2.** Bioscore comparison plots. Each point represents a unique sampling site. The black line is the fitted curve from a model II regression using a major axis regression and the gray line a 1:1 line. Panel A is for the ASPT (macroinvertebrates) and Panel B is for the TDI (diatom) metrics.

associated with HQA for thresholds of 59 to 63% (Fig. 5). Transitivity ranged from 0.1 to 1 across all subnetworks and threshold with mean transitivity across the thresholds being 0.791 (SD = 0.146). Linkage density was not associated with HQA, but did correlate significantly (negatively) with HMS between thresholds of 24–66%. Linkage density ranged from 0.07 to 1 across all subnetworks and thresholds with means across the thresholds was 0.554 (SD = 0.315). Cohesion was significantly ($p < 0.05$) negatively associated with HMS sporadically between thresholds of 7% to 79% and was not associated with HQA for any threshold values. Cohesion ranged from 0 to 144 with a mean of 19.08 (SD = 30.28).

## 4. Discussion

We have shown that eDNA based biomonitoring reveals similar trends at a national scale to those produced via traditional methods for benthic macroinvertebrate and diatom biomonitoring metrics. We also show that 18S-derived biodiversity holds greater potential as an eDNA biomonitoring marker compared to COI or rbcL and that eDNA in general captured more biodiversity compared to traditional sampling methods. Lastly, we show that multiple levels of eDNA-based biomonitoring data can be screened for environmental sensitivity and, when combined with co-occurrence networks, distill additional community based environmental assessment that incorporates environmental
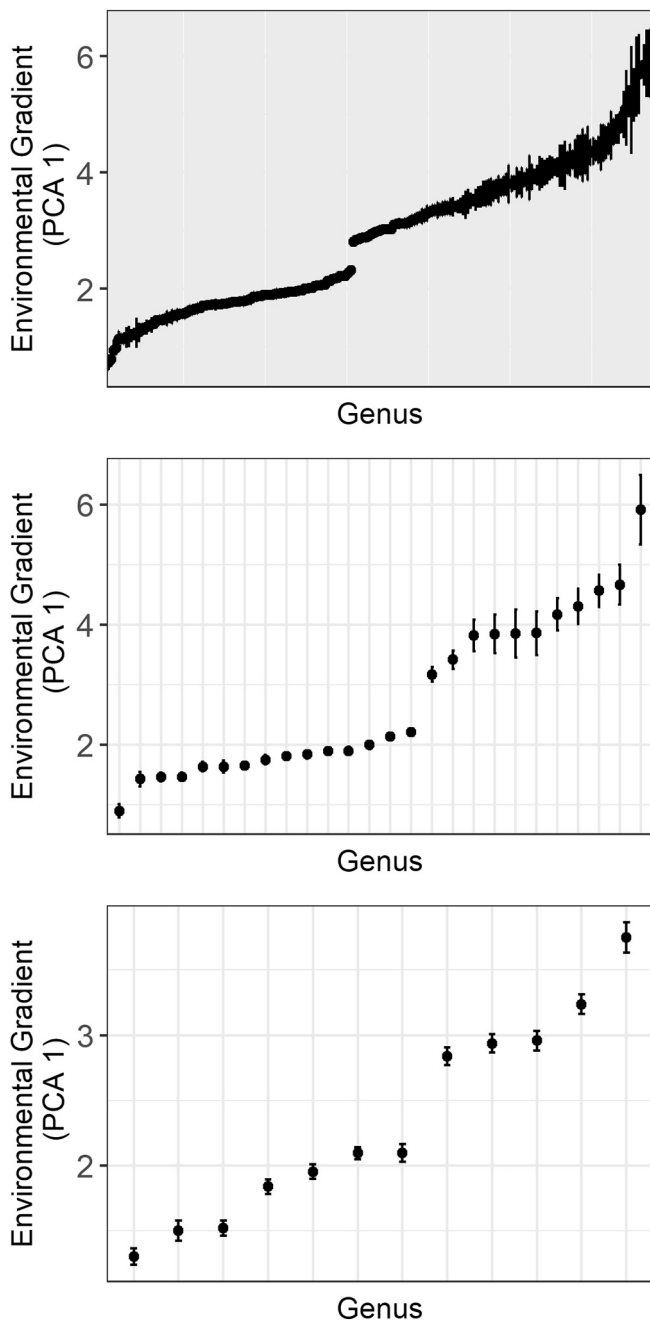
**Fig. 3.** Environmental DNA derived indicator taxa (genera) utilizing universal 18S primers (top), COI (middle), rbcL (bottom). Each panel shows all Genera (x-axis) and their environmental standard deviation (y-axis).

filtering and indicator taxa interactions, which will undoubtedly impact future regulatory protocols.

Traditional biomonitoring metric scores such as WHPT and BMWP, from which ASPT are derived, and TDI are based on indicator taxa likelihood occurrence for sites along pollution or stressor gradients. The exact measures used for benthic macroinvertebrates are rather cryptic in the corresponding government reports, but have been distilled to relate to testing of macroinvertebrate group persistence for two sites for the UK and the broader EU (Hawkes, 1998; Water Framework Directive, 2014b). Diatom biometric scores are more transparent and are based on individual species occurrences across a set of environmental conditions using weighted averages; however the individual metric scores are not accessible in the corresponding report and need to be

obtained as part of a software program called DARLEQ2 (Kelly et al., 2008). Comparisons between eDNA and traditional metrics show positive covariation between approaches for both diatoms and invertebrates, with diatom metrics calculated from eDNA data showing higher values compared to values calculated using traditional methods, whereas the macroinvertebrate eDNA metric was slightly lower than the traditional metric. Taxa assignment overlap between eDNA and traditional methods, and was 47.83% for macroinvertebrate indicator groups versus 15.57% for genera, with eDNA generally yielding higher diversity than traditional methods for each set of metrics, though not for all sites. The disparity in overlap between indicator group and genera between eDNA and traditional assignment is largely due to the mixed class identification that is utilized for traditional invertebrate biomonitoring, and is often limited to family level (Morse et al., 2007), therefore many indicator taxa are not identified to the genus level. Diatom overlap was 11% for indicator groups versus 33% for genera, which may reflect the extensive disparity of diatom taxonomy names between existing diatom studies and databases (Mann, 1999). Of note was the inability of eDNA to adequately sample macroinvertebrates from 2 sites, which led to poor eDNA assessment. As such, the poor eDNA taxonomic coverage was not noticeable until after sequencing, while a poor kick net sample can be identified during the field sampling itself. Overall, the few studies that have assessed eDNA biomonitoring, including the present study, show that eDNA based methods generally produce higher diversity, but do vary in the taxonomic groups they identify, which can be linked to disparity in the sampling methods used (Gibson et al., 2015; Elbrecht et al., 2017; Hajibabaei et al., 2019).

Taxonomic annotation for any method (e.g. traditional or eDNA) is inherently linked to the sequence database or identification key used to assign the final taxonomic designation, and is expected to fall short of fully cataloging the biodiversity of any given site. For traditional kick-net sampling, taxonomic assignment is often limited to family or genera level assignments and often omits/simplifies small or difficult to distinguish groups that are key environmental indicators, such as Rotifers, Oligochaeta or Chironomidae (Furse et al., 2009; Elbrecht et al., 2017). With regards to efficiency, standard kick-net sampling is estimated to return between 40 and 60% of the expected biodiversity for a given sample (Bradley and Ormerod, 2002; Furse et al., 2009), with an estimated 30% routinely misidentified due to human error (Haase et al., 2010). For metabarcoding based taxonomic assignment, taxonomic identification is linked to the utilized marker and the reference database. This study utilized four unique markers and three unique databased, COI and 12S utilized NCBI, whereas 18S relied on Silva and rbcL utilized the Diat.barcode database. Currently, the coverage of the existing databases are estimated at 14.6% for diatoms (rbcL) 64.5% for freshwater macroinvertebrates (COI), 87.9% for freshwater fish (12S) (Weigand et al., 2019) and 40% for prokaryotes (18S) (Louca et al., 2019). Taxonomic assignment for high throughput sequence data is expected to improve with the ongoing research efforts steadily improving sequence database catalogues, with current assignments still being highly valuable for assessing general ecological trends, including among site environmental assessment (Bista et al., 2017; Cordier et al., 2017).

Environmental filtering/selection of genera (Fig. 2) for the networks found 214 18S genera, associated with environmental variation across the sampling sites, compared to 26 COI (macroinvertebrate) and 12 rbcL (diatom) genera, with no environmental filtering present across the 12S dataset (fish). The reasons behind the disparity in observed indicator groups partially stems from the increased diversity in the 18S derived data whereby 29,948 unique genera were detected across all sites, compared to 171, 138 and 9 for COI, rbcL and 12S respectively. Additionally, the increased phylogenetic depth represented in the 18S derived data allows for increased ecological and evolutionary breadth, potentially making this a better marker for eDNA-based biomonitoring. The finding that 18S provides a better eDNA based biomonitoring marker is in-line with recent findings of other highly diverse molecular
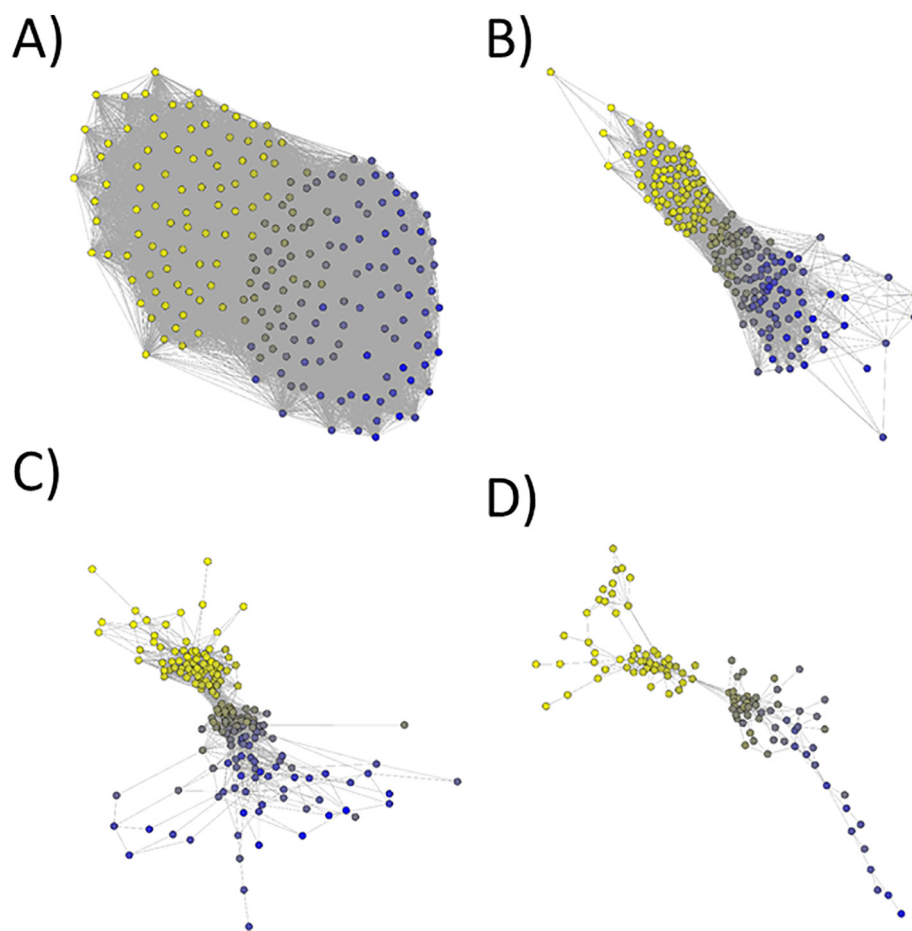
**Fig. 4.** Subset of the main co-occurrence networks constructed. Each node is a unique genus and the color represents the mean gradient value for the node (genera) across the study sites (yellow = low, blue = high). Here we show four of the constructed networks ($N = 157$) for A) 10% co-occurrence B) 30% C) 50% D) 70%.

markers, including 16S, which offer a more powerful assessment of communities across environmental conditions (Pawlowski et al., 2016). Additionally, the COI marker, while widely used for eDNA studies, is not optimal for eDNA based identification likely due to the loss of reads through sequencing of non-target taxa (Collins et al., 2019).

Our co-occurrence analyses found strong associations between genera that co-occurred across sites as descriptors of environmental variation across the network. The high connectivity and degree values associated with a few key genera within the network also highlights the importance of observing not only those taxa associated with a particular environment, but the expected associations of the taxa to other indicator taxa as a means to determine whether an environment is experiencing environmental change (Germain et al., 2018). The network properties associated with community complexity and network connectivity were significantly positively related with habitat modification score (HMS), which relates to highly modified sites having subnetworks that are more centralized in the network, thereby increasing their connectivity properties. This is most evident in the transitivity scores, with the probability that the connecting genera in a network are also connected ecologically. Likewise, the greater linkage density, which is the proportion of realized vs. potential links in the subnetworks is intrinsically higher at more centralized subnetworks (Kantarci and Labatut, 2013). Both these measures highlight the importance of unmodified habitat in protecting not only unique species, but unique communities (Muotka et al., 2002). Additionally, both co-occurrence analyses highlight the potential importance of key individual indicator taxa within their respective networks via cohesion, which indicates the number of vertices that can be removed prior to the network

collapsing (White and Harary, 2001). The resulting networks also provide information regarding which key indicator interactions could provide a better indication of change in environmental condition of a site, rather than only presence/absence of the indicator taxa themselves, which may occur stochastically across a range of environmental conditions (Beauchard et al., 2017).

Overall, this study shows the application of eDNA metabarcoding as a viable surrogate for traditional biomonitoring methods and also shows the ability to incorporate a wide range of taxa that have been overlooked using traditional sampling and biomonitoring protocols. We emphasize the need for future studies to evaluate communities for co-occurrence to ensure that false positives from random associations between indicators and local random/stochastic factors are not associated with co-occurring indicator groups. This approach should provide a more sensitive assessment of community response to environmental variation.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.138801.

## CRediT authorship contribution statement

**Mathew Seymour:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft. **Francois K Edwards:** Conceptualization, Methodology, Investigation, Writing - review & editing. **Bernard J Cosby:** Conceptualization, Methodology. **Martyn G Kelly:** Writing - review & editing. **Mark de Bruyn:** Conceptualization, Writing - review & editing. **Gary R Carvalho:** Writing - review & editing, Conceptualization. **Simon Creer:** Conceptualization.
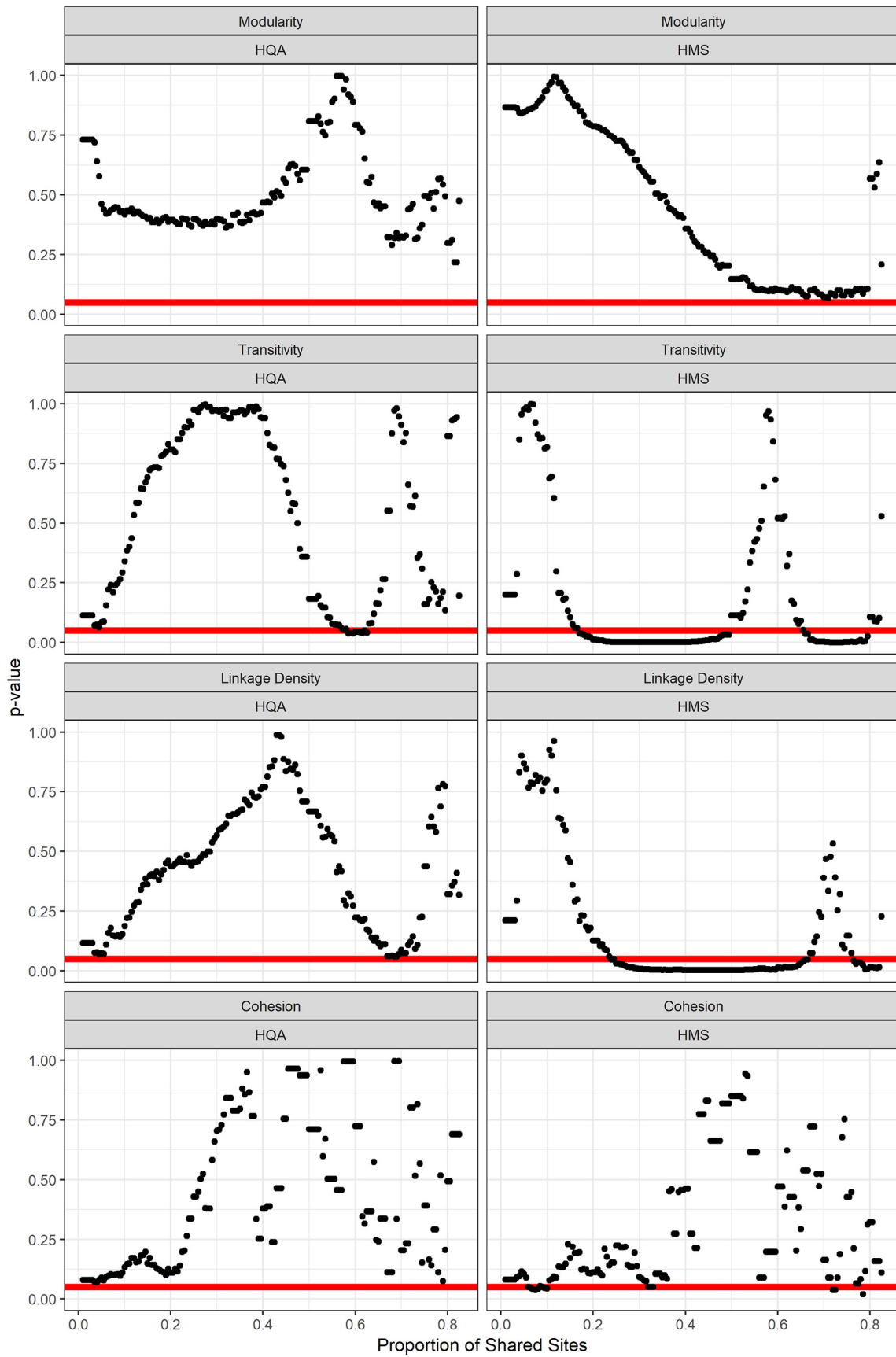
**Fig. 5.** Sensitivity analyses of the co-occurrence network threshold (x-axis) and their associated subnetworks (i.e. sampling site) properties (top: modularity, 2nd row: transitivity, 3rd row: linkage density, bottom: cohesion) tested against HQA (first column) and HMS (second column). The red line indicates the 0.05 significance value of the x-axis (p-value for the respective test). Points on or below the line signify a significant association between the network property and the ecological assessment metric (HQA or HMS).

## Declaration of competing interest

None.

## Acknowledgements

## Data availability

The data associated with the manucript is available on Mendeley Data (DOI: 10.17632/zcfywkyr7j.2)

## References

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

Araújo, M.B., Rozenfeld, A., Rahbek, C., Marquet, P.A., 2011. Using species co-occurrence networks to assess the impacts of climate change. Ecography 34 (6), 897–908. https://doi.org/10.1111/j.1600-0587.2011.06919.x.

Baird, D.J., Hajibabaei, M., 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Mol. Ecol. 21 (8), 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x.

Beauchard, O., Veríssimo, H., Queirós, A.M., Herman, P.M.J., 2017. The use of multiple biological traits in marine community ecology and its potential in ecological indicator development. Ecol. Indic. 76, 81–96. https://doi.org/10.1016/j.ecolind.2017.01.011.

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

Bista, I., Carvalho, G.R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Creer, S., 2017. Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nat. Commun. 8, 14087. https://doi.org/10.1038/ncomms14087.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., ... de Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology &amp; Evolution 29 (6), 358–367. https://doi.org/10.1016/j.tree.2014.04.003.

Bradley, D.C., Ormerod, S.J., 2002. Evaluating the precision of kick-sampling in upland streams for assessments of long-term change: the effects of sampling effort, habitat and rarity fig: 5 tab: 5. Arch. Hydrobiol. 199–221.

Cadotte, M.W., Tucker, C.M., 2017. Should environmental filtering be abandoned? Trends Ecol. Evol. 32 (6), 429–437. https://doi.org/10.1016/j.tree.2017.03.004.

Carvalho, L., Mackay, E.B., Cardoso, A.C., Baattrup-Pedersen, A., Birk, S., Blackstock, K.L., ... Solheim, A.L., 2019. Protecting and restoring Europe's waters: An analysis of the future development needs of the Water Framework Directive. Science of The Total Environment 658, 1228–1238. https://doi.org/10.1016/j.scitotenv.2018.12.255.

Collins, R.A., Bakker, J., Wangensteen, O.S., Soto, A.Z., Corrigan, L., Sims, D.W., ... Mariani, S., 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution 0 (0). https://doi.org/10.1111/2041-210X.13276.

Cordier, T., Esling, P., Lejzerowicz, F., Visco, J., Ouadahi, A., Martins, C., Pawlowski, J., 2017. Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. Environmental Science & Technology 51 (16), 9118–9126. https://doi.org/10.1021/acs.est.7b01518.

Cox, M.P., Peterson, D.A., Biggs, P.J., 2010. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11 (1), 485.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695 (5), 1–9.

Edgar, R.C., 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. BioRxiv, 81257 https://doi.org/10.1101/081257.

Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J., Leese, F., 2017. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods Ecol. Evol. 8 (10), 1265–1275. https://doi.org/10.1111/2041-210X.12789.

Emmett, B.E., Abdalla, M., Anthony, S., Astbury, S., August, T., Barrett, G., ... Bradley, D., 2017. Glastir Monitoring & Evaluation Programme. Final report.

Furse, M.T., Hering, D., Brabec, K., Buffagni, A., Sandin, L., Verdonschot, P.F.M., 2009. The Ecological Status of European Rivers: Evaluation and Intercalibration of Assessment Methods. 188. Springer Science & Business Media.

Gavrilescu, M., Demnerová, K., Aamand, J., Agathos, S., Fava, F., 2015. Emerging pollutants in the environment: present and future challenges in biomonitoring, ecological risks and bioremediation. New Biotechnol. 32 (1), 147–156. https://doi.org/10.1016/j.nbt.2014.01.001.

Germain, R.M., Mayfield, M.M., Gilbert, B., 2018. The 'filtering' metaphor revisited: competition and environment jointly structure invasibility and coexistence. Biol. Lett. 14 (8), 20180460. https://doi.org/10.1098/rsbl.2018.0460.

Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I., Hajibabaei, M., 2015. Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLoS One 10 (10), e0138432 Retrieved from. https://doi.org/10.1371/journal.pone.0138432.

Glover, R., 2019. Biomonitoring and Surveillance with Short-and Long-Read Metabarcoding.

Haase, P., Pauls, S.U., Schindehütte, K., Sundermann, A., 2010. First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. J. N. Am. Benthol. Soc. 29 (4), 1279–1291. https://doi.org/10.1899/09-183.1.

Hajibabaei, M., Porter, T.M., Robinson, C.V., Baird, D.J., Shokralla, S., Wright, M.T.G., 2019. Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. PLoS One 14 (12), e0225409 Retrieved from. https://doi.org/10.1371/journal.pone.0225409.

Hawkes, H.A., 1998. Origin and development of the biological monitoring working party score system. Water Res. 32 (3), 964–968. https://doi.org/10.1016/S0043-1354(97)00275-3.

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

Kantarci, B., Labatut, V., 2013. Classification of complex networks based on topological properties. 2013 International Conference on Cloud and Green Computing, pp. 297–304. https://doi.org/10.1109/CGC.2013.54.

Karimi, B., Maron, P.A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D., Ranjard, L., 2017. Microbial diversity and ecological networks as indicators of environmental quality. Environ. Chem. Lett. 15 (2), 265–281. https://doi.org/10.1007/s10311-017-0614-6.

Kelly, M., 1998. Use of the trophic diatom index to monitor eutrophication in rivers. Water Res. 32 (1), 236–242. https://doi.org/10.1016/S0043-1354(97)00157-7.

Kelly, M., Juggins, S., Guthrie, R., Pritchard, S., Jamieson, J., Rippey, B., Yallop, M., 2008. Assessment of ecological status in U.K. rivers using diatoms. Freshw. Biol. 53 (2), 403–422. https://doi.org/10.1111/j.1365-2427.2007.01903.x.

King, G., Zeng, L., 2001. Logistic regression in rare events data. Polit. Anal. 9 (2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868.

Lampert, A., 2019. Over-exploitation of natural resources is followed by inevitable declines in economic growth and discount rate. Nat. Commun. 10 (1), 1419. https://doi.org/10.1038/s41467-019-09246-2.

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front. Zool. 10 (1), 34. https://doi.org/10.1186/1742-9994-10-34.

Louca, S., Mazel, F., Doebeli, M., Parfrey, L.W., 2019. A census-based estimate of Earth's bacterial and archaeal diversity. PLoS Biol. 17 (2), e3000106. https://doi.org/10.1371/journal.pbio.3000106.

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

Mandaville, S.M., 2002. Benthic Macroinvertebrates in Freshwaters: Taxa Tolerance Values, Metrics, and Protocols. Soil & Water Conservation Society of Metro Halifax, Halifax, Canada Retrieved from. http://lakes.chebucto.org/H-1/tolerance.pdf.

Mann, D.G., 1999. The species concept in diatoms. Phycologia 38 (6), 437–495. https://doi.org/10.2216/i0031-8884-38-6-437.1.

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. Open Science, 2(7). Retrieved from http://rsos.royalsocietypublishing.org/content/2/7/150088.abstract.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., ... Iwasaki, W., 2020. 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.

Morse, J.C., Bae, Y.J., Munkhjargal, G., Sangpradub, N., Tanida, K., Vshivkova, T.S., ... Yule, C.M., 2007. Freshwater biomonitoring with macroinvertebrates in East Asia. Frontiers in Ecology and the Environment 5 (1), 33–42. https://doi.org/10.1890/1540-9295(2007)5[33:FBWMIE]2.0.CO;2.

Muotka, T., Paavola, R., Haapala, A., Novikmec, M., Laasonen, P., 2002. Long-term recovery of stream habitat structure and benthic invertebrate communities from in-stream restoration. Biol. Conserv. 105 (2), 243–253. https://doi.org/10.1016/S0006-3207(01)00202-6.

Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., Esling, P., 2016. Protist metabarcoding and environmental biomonitoring: time for change. Eur. J. Protistol. 55, 12–25. https://doi.org/10.1016/j.ejop.2016.02.003.

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.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-

based tools. Nucleic Acids Res. 41 (D1), D590–D596. https://doi.org/10.1093/nar/gks1219.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna Retrieved from. http://www.r-project.org.

Rimet, F., Gusev, E., Kahlert, M., Kelly, M.G., Kulikovskiy, M., Maltsev, Y., ... Bouchez, A., 2019. Diat.barcode, an open-access curated barcode library for diatoms. Scientific Reports 9 (1), 15116. https://doi.org/10.1038/s41598-019-51500-6.

Schmieder, R., Lim, Y.W., Rohwer, F., Edwards, R., 2010. TagCleaner: identification and removal of tag sequences from genomic and metagenomic datasets. BMC Bioinformatics 11 (1), 341.

Seymour, M., 2019. Rapid progression and future of environmental DNA research. Communications Biology 2 (1), 80. https://doi.org/10.1038/s42003-019-0330-9.

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

Simboura, N., Panayotidis, P., Papathanassiou, E., 2005. A synthesis of the biological quality elements for the implementation of the European Water Framework Directive in the Mediterranean ecoregion: the case of Saronikos Gulf. Ecol. Indic. 5 (3), 253–266. https://doi.org/10.1016/j.ecolind.2005.03.006.

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

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol. Ecol. 19 (s1), 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x.

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Domaizon, I., 2018. Avoiding quantification bias in metabarcoding: application of a cell biovolume correction factor in diatom molecular biomonitoring. Methods Ecol. Evol. 9 (4), 1060–1069. https://doi.org/10.1111/2041-210X.12960.

Water Framework Directive, 2014a. UKTAG River Assessment Method Macrophytes and Phytobenthos.

Water Framework Directive, 2014b. Water Framework Directive, United Kingdom Advisory Group, 2014. Invertebrates (General Degradation) Whalley, Hawkes, Paisley & Trigg (WHPT) Metric in River Invertebrate Classification Tool (RICT), UKTAG River Assessment Method. Benthic Invertebrate Fauna ReportUKTAG.

Weigand, H., Beermann, A.J., Čiampor, F., Costa, F.O., Csabai, Z., Duarte, S., Ekrem, T., 2019. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. Sci. Total Environ. 678, 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247.

White, D.R., Harary, F., 2001. The cohesiveness of blocks in social networks: node connectivity and conditional density. Sociol. Methodol. 31 (1), 305–359. https://doi.org/10.1111/0081-1750.00098.