GCN sample analysis via single species assay and metabarcoding

Extraction and analysis of Great Crested Newt eDNA samples via single species and metabarcoding

April 2024

Natural England Commissioned Report NECR534

www.gov.uk/natural-england

NATURAL ENGLAND

About Natural England

Natural England is here to secure a healthy natural environment for people to enjoy, where wildlife is protected and England's traditional landscapes are safeguarded for future generations.

Further Information

This report can be downloaded from the Natural England Access to Evidence Catalogue. For information on Natural England publications or if you require an alternative format, please contact the Natural England Enquiry Service on 0300 060 3900 or email enquiries@naturalengland.org.uk.

Copyright

Catalogue code: NECR534

Report details

Author(s)

Helen C. Rees, Steven D., Kane

Natural England Project Manager

Debbie Leatherland Debbie.Leatherland@naturalengland.org.uk

Contractor

RSK ADAS Ltd.

Keywords

Amphibian, eDNA, freshwater, Great crested newt, qPCR, *Triturus cristatus*.

Acknowledgements

We would like to thank Debbie Leatherland of Natural England for commissioning the project; Cellmark for sending the samples to our laboratories; and Nick Dunn and Debbie Leatherland for their assistance during the project.

Citation

This report should be cited as: REES, H.C., KANE, S.D. GCN sample analysis via single species assay and metabarcoding, 2024. NECR534. Natural England.

Foreword

This work was commissioned to build on a recent evidence review of great crested newt (GCN) eDNA monitoring protocols (NECR476) and, on a small scale, compare results from two eDNA capture methods (ethanol precipitation and filtration) for the detection of the species. The work here also seeks to identify whether metabarcoding is comparable with single-species detection of GCN. The results from this project will contribute to an evaluation of the current methodology for monitoring GCN, and help support a shift towards a revised methodology, if the need for an updated protocol is identified.

Natural England commission a range of reports from external contractors to provide evidence and advice to assist us in delivering our duties. The views in this report are those of the authors and do not necessarily represent those of Natural England.

Executive summary

Following the recent review and publication of: An evidence review for great crested newt eDNA monitoring protocols - NECR476 (naturalengland.org.uk), Natural England would like to increase the evidence base in the area of eDNA sampling (ethanol precipitation vs filtration). The aim of this project was to compare the results of great crested newt (GCN) environmental DNA (eDNA) analysis from samples collected using either ethanol precipitation or filtration methods. The analysis of samples was performed using the single species GCN assay following laboratory protocols in: Analytical and methodological development for improved surveillance of the Great Crested Newt, and other pond vertebrates - WC1067. Furthermore, Natural England wished to investigate whether GCN could be reliably detected from a sub-set of the samples collected via filtration methods using DNA metabarcoding.

To do this, 100 samples (including field blanks) were collected using the standard ethanol precipitation-based sampling kits (as per WC1067) and during the same sampling visit a sample was also taken using a Sterivex-HV Pressure Filter Unit with a 0.45µm pore size, and the addition of an ethanol based preservative solution. The ethanol precipitation samples were analysed by Cellmark (a GCN eDNA analysis service provider) using the methods detailed in WC1067 and the results were shared for this report to allow comparison with the results of the filter-based sample results. After DNA was extracted from the filter samples, GCN analysis was performed using the methods in WC1067 and via metabarcoding of a sub-set of 20 of the samples all of which were positive for GCN with both sampling methods.

GCN species-specific results comparison of the two sample collection methods were largely in agreement with 27 of the samples being positive for GCN eDNA by both methods with 5 additional positive samples for both the ethanol precipitation and filtration collected samples respectively. Metabarcoding of 12S DNA for vertebrate species resulted in 29 species being detected across the 20 samples with 11 of the 20 samples detecting great crested newt DNA.

Contents

Introduction

Natural England is the Government's advisor for the natural environment. It provides practical advice on how to safeguard England's natural wealth for the benefit of everyone. RSK ADAS is an environmental consultancy which exists to provide ideas, specialist knowledge and solutions to secure our food and enhance the environment.

Natural England wishes to compare the effectiveness of ethanol precipitation and filtration eDNA capture methods for use in species-specific GCN rt-PCR analysis. The current methodology for collection of water samples for analysis for GCN eDNA uses ethanol precipitation of eDNA, one of the first methods described for eDNA recovery (Ficetola and others. 2008, Biggs and others. 2014). Its use was adopted in 2014 within the Natural England technical advice note (WC1067) after extensive validation of a PCR assay that was initially described by Thomsen and others. (2012) for GCN detection, in conjunction with DNA collection methods based on earlier works by Ficetola and others. (2008). Since then, there have been extensive development and use of filter based eDNA capture for the sampling of eDNA from waterbodies (Turner and others. 2014, Wilcox and others. 2015, Hosler 2017, Deiner and others. 2018, Sepulveeda and others. 2019 etc.) and as such filtration has now largely replaced ethanol precipitation as the method of choice for capturing eDNA from water samples.

Although currently only 90 mL of water are sampled from a pond, the results generated for GCN eDNA are well tuned to ecological survey in that when GCN are present by ecological survey they can be detected by eDNA and vice versa. Any perceived increases in sensitivity afforded by extraction from the potentially greater sample volumes obtained by using filtration-based sample collection could undermine previous survey efforts by detecting low/negligible GCN presence in ponds that were previously negative; and increase the detections of eDNA from those low transient populations of GCN that are not picked up by traditional survey. Regardless of this, the methods used in the Natural England technical advice note, despite being the most highly validated eDNA assay currently being used in a regulatory framework (Thalinger and others. 2021) now appear out-of-step with current eDNA methods and it is important to provide new evidence to allow Natural England to fully review current sampling methodologies.

Aims and Objectives

The overall aim of the project was to increase the evidence base to allow the further review of great crested newt eDNA monitoring protocols. The main aim was to compare the use of the current ethanol precipitation-based sampling methodology with filtration-based sampling in terms of the results of the species-specific GCN assay as described in WC1067. Additionally, a further aim was to compare the reliability of metabarcoding for the detection of GCN from the filtration-based samples.

This report details the methodology employed in these studies, the results obtained and, discussion of the results. All data will be made available for further study and could be used for a training day for Natural England staff on the DNA approaches used.

Materials and Methods

Sample collection

Three Natural England District Level Licensing (DLL) scheme areas were chosen to collect filtration samples in 2023, alongside standard precipitation surveys. Sample collection was completed by Habitat Delivery Bodies (HDBs) based in Cheshire (1 HDB each covering Cheshire East, and Cheshire West), Norfolk (1 HDB) and Yorkshire (3 HDBs) (Table 1).

A subset of ponds within each area was chosen for sampling. This was a semi-random selection from ponds created or restored prior to 1st April 2022, ensuring a spread of ponds within each scheme area across the following criteria:

- Geographic spread across the scheme area based on the pond's district.
- Age of pond (ranging from the first to the fourth year of monitoring).
- Whether the pond was a creation or restoration.
- Previous monitoring results via precipitation eDNA testing (if carried out) – present, absent, or inconclusive (and if inconclusive, whether this was due to inhibition or degradation).
- Whether the pond is within a 'core', 'fringe' or 'outside' the modelled DLL Strategic Opportunity Areas for GCN

Samples were taken using 0.45 µm sterivex filter kits prepared and supplied by Surescreen Scientific (under a separate contract). Those taking samples were also able to access a video made by Surescreen Scientific on how to take samples via filtration, in addition to written guidance. HDBs were advised to use bottled water for their control samples. For ease, HDBs returned their filtration samples to the same address as precipitation samples (Orchid Cellmark), and these were stored at room temperature for up to eight weeks before shipping to RSK ADAS.

Table 1. Number of samples collected in each area.

Region	Planned Number of Samples by Area
Cheshire East	15 ponds; 2 controls
Cheshire West	15 ponds; 2 controls
Norfolk	26 ponds; 2 controls
Yorkshire	4 ponds; 2 controls

Region	Planned Number of Samples by Area
Yorkshire	12 ponds; 2 controls
Yorkshire	16 ponds; 2 controls
Total	88 ponds; 12 controls

Laboratory standard and specifications

All laboratory activities associated with DNA analysis are subject to errors if quality control is inadequate. Our DNA analysis follows a unidirectional workflow with separate laboratories and staff to act as a physical separation for the different aspects of the analysis work. This greatly reduces the potential for contamination of samples or the PCR amplicons. 'Blank' PCRs (sterile water rather than DNA) are used to monitor for reagent/procedural contamination, and in addition positive control samples are used to increase confidence in the results and identify any cross-contamination issues, should they occur.

DNA extraction from filters

DNA was extracted from all filters using the DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions with minor modifications (below) and finally resuspended in 200 µL of elution buffer (Appendix 1). Extraction blanks were included (1 per day of DNA extractions) to monitor for any cross-contamination during the DNA extraction.

- 540 µL ATL buffer, 300 µL AL buffer, and 50 µL proteinase K (all components of the kit) used instead of standard amounts.
- Addition of $5.7x10^{-5}$ µg/µL synthetic control DNA for DNA extraction efficiency monitoring
- 500 µL 100% ethanol used instead of standard amount.

All DNA extracts were quantified using a Qubit 3.0 Fluorometer (Invitrogen) following the manufacturer's instructions then stored at -20 ºC prior to PCR set up (Appendix 1).

The efficiency of DNA extraction was measured by the addition of known concentration of a synthetic DNA control to the buffers added to the filter. The recovery of this synthetic DNA was measured by a qPCR assay specific for the sequence of the synthetic DNA control (Appendix 1) and was used as a proxy for total eDNA extraction efficiency with acceptable limits being within 2 standard deviations of the average Cq value (95% of samples should be within this range).

eDNA assay

GCN eDNA assay was performed in accordance with WC1067 which amplifies an 81 bp fragment of the cytochrome b gene (Table 2, Appendix 1). All samples were subjected to testing for eDNA recovery and inhibition.

Table 2. GCN primer and probe sequences (Thomsen and others. 2012).

Oligo Name	Sequence (5' - 3')
TCCBF	CGTAAACTACGGCTGACTAGTACGAA
TCCBR	CCGATGTGTATGTAGATGCAAACA
TCCB.probe	FAM-CATCCACGCTAACGGAGCCTCGC-BHQ1

Metabarcoding PCR

The primer combination used for the first round PCR amplification was 12S forward and reverse for 12S PCRs (Riaz and others. 2011). Overhang adapter sequences (Table 3) were included at the 5' end of the primers to ensure compatibility with Illumina index and sequencing adapters (Illumina 2012). PCRs included one negative control (ddH_2O in place of DNA); two DNA extraction blanks; a positive control sample (*Scomber scombrus* Atlantic Mackerel); and all pond DNA samples.

Table 3. Primers used for metabarcoding first round PCR.

Primer Name	Oligonucleotides (5'-3')	Reference
12S forward (plus adapter)	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**TAGAACAGGCTCCTCTAG	Riaz (2011)
12S reverse (plus adapter)	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**TTAGATACCCCACTATGC	Riaz (2011)
Index 1	CAAGCAGAAGACGGCATACGAGATXXXXXXXX**GTCTCGTGGGCTCGG**	Illumina (2011)
Index 2	AATGATACGGCGACCACCGAGATCTACACXXXXXXXX**TCGTCGGCAGCGTC**	Illumina (2011)

Note 1: For the 12S primers, sequences marked in bold are the first round PCR primer adapter sequences, the remainder are the 12S locus-specific primer sequences. For the Index primers, sequences marked in bold are Illumina overhang adapter sequences, Index 1 and 2 sequences are marked with Xs as this sequence is variable for each different sample, those in normal text are the P5 and P7 sequences. Index 1 (i7) and Index 2 (i5) are examples of the type of primers used with the Index sequence itself being altered for different samples.

The first round PCR amplicons for each sample were pooled and run on a 1.5% agarose gel. Any bands of the correct size were excised and purified using NucleoSpin® Gel and PCR Clean-up purification columns (Macherey-Nagel) according to the manufacturers' instructions (Appendix 1).

Sequence library preparation and sequencing

The second round of PCR or 'Index' PCR was performed by Source Bioscience to add molecular identification (MID) tags (unique 8-nucleotide sequences) and Illumina MiSeq sequencing adapters to the first round PCR products.

The indexed amplicons were quantified via a fluorometric method involving QuantiFluor dsDNA assay (Promega); and qualified using electrophoretic separation on the Agilent Fragment Analyzer 5300. This concentration and sizing information has been used to calculate the molarity of each sample. All samples passed QC checks carried out by Source Bioscience. Libraries were then pooled in equimolar amounts to create one library for Illumina sequencing. The amplicon library pool was spiked with 20 % PhiX Control v3 library (Illumina) and run on the Illumina MiSeq using a MiSeq Reagent Kit v2 500 cycle kit (Illumina), to generate 250-bp paired-end reads. PhiX DNA is derived from the small, well characterized bacteriophage PhiX genome, it is a concentrated Illumina library (10 nM in 10 µl) that has an average size of 500 bp and consists of balanced base composition at ~45% GC and ~55% AT and serves as an in-run QC for the Illumina sequencing.

Bioinformatic processing

Data processing was performed on an Intel i7 PC running Ubuntu Linux 20.04.4 LTS. In an initial step, paired end reads were trimmed using trimmomatic 0.39 (Bolger, Lohse and Usadel, 2014) to remove Miseq adapters, to clip low quality and unpaired reads, and to truncate the sequence if the average phred score of a 5nt sliding window dropped below 25. An example command line to process a demultiplexed FASTQ sequence file was:

java -jar trimmomatic-0.39.jar PE R1_001.fastq R2_001.fastq read1_paired_R1_001.fastq read1_unpaired_R1_001.fastq read2_paired_R2_001.fastq read2_unpaired_R2_001.fastq ILLUMINACLIP:NexteraPE-PE.fa:2:30:10:2:True LEADING:3 TRAILING:3 MINLEN:36 SLIDINGWINDOW:5:25

Paired reads were merged using FLASH 1.2.11 (Fast Length Adjustment of SHort reads, Magnoc and Salzberg, 2011) to convert paired end reads (R1 and R2 in the MiSeq platform) to a single merged read using a minimum overlap of 80 nucleotides and a maximum of 150 nucleotides.

After converting DNA sequences from FASTQ format to FASTA format using SeqKit (Shen and others, 2016), template specific PCR primers at the 5' and 3' ends were removed using the "linked adapter" option of Cutadapt 3.5 (Martin, 2011) with a 10% error rate

within the primer site i.e. 2 bp variants allowed per primer. Only trimmed sequences i.e. those containing both matching primer sequences were retained.

Before taxonomic assignment, standard Linux tools were used to identify 100% identical reads and condense them down to a single read to minimise time-consuming and repetitive BLAST searches, However, a record of the frequency of replicate sequences was maintained. Any reads with less than 5 replicates were excluded from the BLAST search as these were considered sequencing errors and omitted from further analyses (Harper and others 2018).

A custom 12S BLAST database was created on 09/10/2023 from the National Centre for Biotechnology Information (NCBI) database using the search terms ‘vertebrata’ AND ‘12S’ before downloading the records in FASTA format. A total of 190K sequences downloaded from NCBI were included in the final database.

BLAST searching was performed using the “megablast” program which is optimised to identify alignments in highly similar sequences and returned the top hit for each query sequence in a custom tabulated format. An e-value of 1e-15 was set; higher values such as 1 or 10 return a larger list of more low-scoring hits, and actual e-values returned were in the order of 1e-150 for a full-length alignment.

A custom perl script filtered the BLAST output, identifying hits sharing an accession number and passing a set of criteria covering the percentage similarity between the query sequence and the database sequence (≥97%), and having a query alignment length difference less than 6 bp. Read counts for each sequence passing the similarity and query alignment length filters were pooled based on accession number to generate a final frequency count for each accession.

Results

DNA extraction

DNA was extracted from the 92 individual filter samples returned to RSK ADAS Ltd. and the DNA quantified. Volumes of water sampled by the Habitat delivery bodies and the volume of recovered preservative were also recorded (Table 4). Return of the filters did not raise any issues with leakage of preservation solution, however six of the samples had lower than 1 mL preservation solution recovered indicating that on occasion not enough preservation solution was added during samples collection (these are noted in Table 4). The efficiency of DNA extraction was within acceptable limits for all but four samples. These were not the same samples as those with low volumes of recovered preservative.

Table 4. Filter sample information. Some water volumes were not supplied by the samplers and are marked as 'unknown'. Some samples marked 'Filtration field blank' were not collected via ethanol precipitation and thus there were no corresponding precipitation results for these samples. DNA concentrations after 12S PCR amplicon clean up are only available for the 20 samples chosen for metabarcoding, all others are marked 'N/A'.

Kit ID Number	Sample ID	Volume water filtered (mL)	Volume preservative recovered (mL)	GCN result filtration (positives out of 12)	GCN result ethanol precipitation (positives out of 12)	DNA extract concentration (ng/µL)
FK1170	B2-027	300ml	2.3	12	9	2.08
FK1171[#]	B2-030	400ml	0.8	0	0	0.406
FK1172	B2-032 (2)	500ml	1.7	0	Filtration field blank	ND
FK1173	B2-030 (2)	400ml	1.6	0	Filtration field blank	ND
FK1174	B2-032	175ml	1.7	0	0	0.057
FK1175	B2-029	300ml	2.3	6	11	0.605
FK1176	S1-004	141ml	2.0	0	0	1.76
FK1177	S1-068	unknown	2.0	0	0	0.879
FK1178	S1-069	500ml	2.4	5	2	3.57
FK1179	S1-070	500ml	1.6	0	0	4.33
FK1180	S1-027	500ml	1.8	0	0	0.463
FK1181	S1-069	500ml	2.0	0	Filtration field blank	0.074
FK1182[#]	C2-029	78ml	0.6	0	1	1.46
FK1183	S1-054	300ml	2.4	0	0	2.4
FK1184	S1-061	89ml	1.7	0	0	0.148
FK1185	S1-010	250ml	1.8	0	1	3.68
FK1186*	C2-034	30ml	1.7	0	0	0.128
FK1187	C2-001	300ml	1.7	0	0	1.22
FK1188*	S1-004	unknown	1.4	0	Filtration field blank	ND
FK1189	S1-070	400ml	1.7	0	Filtration field blank	ND
FK1191	Z1-001	150ml	1.9	12	12	4.25
FK1192	Z3-080	500ml	1.8	0	1	0.801
FK1193	Z3-002	500ml	1.9	1	no result sent	5.82
FK1194	Z1-011	50ml	1.8	0	0	1.56
FK1195	Z1-038	180ml	2.0	11	12	3.78
FK1196	Z1-029	300ml	2.0	11	10	5.74
FK1197	Z1-035	300ml	1.6	0	0	1.71
FK1198	Z1-033	180ml	2.0	0	0	4.38

Kit ID Number	Sample ID	Volume water filtered (mL)	Volume preservative recovered (mL)	GCN result filtration (positives out of 12)	GCN result ethanol precipitation (positives out of 12)	DNA extract concentration (ng/µL)
FK1199	Z1-024	500ml	1.8	0	0	10.9
FK1200	Z1-010	90ml	1.9	0	0	0.215
FK1201	Z1-023	300ml	2.0	1	0	4.21
FK1203	Z1-004	500ml	1.8	3	1	18.1
FK1204	Z1-007	500ml	1.8	0	0	1.61
FK1205	Z1-034	500ml	1.5	12	12	2.35
FK1208	Z3-009	250ml	1.8	0	0	4.96
FK1209	Z1-014	120ml	1.5	12	6	1.11
FK1211	Z1-013	500ml	1.5	0	0	2.25
FK1212	Z1-036	500ml	2.0	3	0	13.3
FK1213	Z1-025	90ml	1.5	2	1	0.178
FK1215	Z1-012	500ml	1.9	0	0	1.91
FK1216	Z3-008	350ml	1.7	0	0	1.55
FK1217	Z1-018	40ml	1.7	0	0	3.32
FK1218	S1-080	180ml	2.0	0	0	5.85
FK1219	S1-039	500ml	1.8	0	0	5.09
FK1220	B2-016	100ml	1.0	0	0	0.229
FK1221	S1-050	400ml	2.0	0	0	2.57
FK1222*	B2-034	500ml	1.8	0	0	ND
FK1224	S1-075	150ml	2.0	3	0	1.13
FK1225	B2-006	500ml	1.7	0	0	0.18
FK1226	S1-022	500ml	1.6	0	0	0.689
FK1227	B2-022	260ml	2.0	0	0	1.58
FK1228	S1-002	450ml	1.8	12	9	1.15
FK1229	S1-002	500ml	1.8	0	Filtration field blank	0.119
FK1230	S1-041	150ml	1.4	0	0	1.15
FK1231	B2-001	500ml	1.5	0	0	0.157
FK1232	B2-017	250ml	1.6	0	0	0.098
FK1233	B2-033	500ml	1.4	0	0	0.325
FK1234	S1-040	100ml	1.7	0	0	0.791
FK1235	S1-087	400ml	1.8	12	12	3.29
FK1236	B1-027	150ml	1.5	12	2	2.09
FK1237	B1-122	50ml	1.8	4	1	0.428
FK1238	B1-189	350ml	1.7	0	0	0.239
FK1239[#]	B1-141	25ml	0.4	0	0	0.766
FK1240	B1-098	300ml	1.8	12	8	4.34
FK1241	B1-015	unknown	1.8	12	11	3.2
FK1242	B1-067	unknown	1.7	0	0	0.826
FK1243	B1-128	200ml	2.0	8	2	1.65

Kit ID Number	Sample ID	Volume water filtered (mL)	Volume preservative recovered (mL)	GCN result filtration (positives out of 12)	GCN result ethanol precipitation (positives out of 12)	DNA extract concentration (ng/µL)
FK1244#	B1-194	200ml	0.3	0	0	10.2
FK1245	B1-263	300ml	2.1	11	9	0.53
FK1246	B1-046	90ml	1.4	12	6	1.39
FK1247	B1-182	500ml	1.9	0	0	0.988
FK1248	B1-223	250ml	1.8	0	0	4.02
FK1250	B1-108	350ml	2.0	1	0	8.13
FK1251	B1-121	170ml	1.9	2	5	1.13
FK1252	B1-044	125ml	1.8	0	0	2.46
FK1253	C1-004	200ml	1.4	8	1	8.75
FK1254#	C1-016	300ml	0.8	0	0	13.2
FK1255#	C1-106	350ml	0.9	0	0	3.62
FK1256	C1-036	unknown	2.1	0	0	0.347
FK1257	C1-071	200ml	2.4	12	2	22.5
FK1258	C1-037	500ml	1.7	12	12	0.653
FK1259	C1-004	500ml	1.9	11	1	5.23
FK1260	C1-027	500ml	1.8	0	0	ND
FK1261	C1-007	500ml	1.9	0	5	ND
FK1262	C1-114	500ml	1.7	6	1	0.832
FK1263	C1-006	500ml	1.8	11	12	9.25
FK1264	C1-011	500ml	1.6	0	0	2.39
FK1265	C1-063	500ml	1.6	0	0	0.844
FK1266	C1-049	500ml	2.0	5	2	0.201
FK1267	C1-012	500ml	2.1	0	0	4.24
FK1268	C1-022	500ml	1.7	9	0	1.77
FK1269*	C1-088	240ml	1.5	0	1	1.28

Note 1: Dark grey shading represents samples that are positive for GCN by both sampling methods, and light grey shading represents samples that are only positive for GCN with one sampling method.

Note 2: ND marks those samples where DNA concentration was too low to measure.

Note 3: *marks those samples which were shown to have a poor DNA extraction efficiency from filters i.e. were outside of acceptable limits.

Note 4: #marks those samples where less than 1 mL of preservation buffer was recovered.

Species-specific GCN results

A total of 27 of the samples were positive for GCN eDNA by both ethanol precipitation and filtration with five and six additional positive samples respectively by only one sampling method (Figure 1A). However, one result was not received for the ethanol precipitation samples for Z2-002 which was positive (1/12) for GCN using filtration. If removed from the results, then there were five additional positive samples by either one of the sampling methods (Figure 1B). Of the five additional positives when sampled using ethanol precipitation all except one sample showed very low positivity (1/12; C1-005 was 5/12). Of the six additional positives when sampled using filtration two had a PCR score of 1/12 and two a PCR score of 3/12. Only two of the additional positives (B1-108 and C1-022) had a high PCR score of 7/12 or 9/12.

All field blanks were negative for GCN and all ten extraction blanks carried out by RSK ADAS were negative for GCN and all samples were within acceptable limits for inhibition. The remaining 59 filter samples were negative for GCN eDNA, however, four of these samples (C2-034, S1-004, B2-034, and C1-088) were found to be well outside of acceptable limits for the DNA recovery control when defining acceptable limits as within 2 standard deviations of the average Cq value (95% of samples should be within this range) i.e. you would expect 5% of samples to be outside this range. When applying a nominal Ct cutoff of 3 Ct above the expected Ct if 100% of DNA was recovered the same four samples were found to be outside of acceptable limits and therefore had poor DNA recovery efficiencies. S1-004 (a field blank) did not contain detectable concentrations of DNA as would be expected and the same was true for B2-034. C2-034 had a very low DNA concentration (0.128 ng/µL), the remaining sample C1-088 had a lower-than-average concentration of DNA (1.28 ng/µL, average of 3.05 ng/µL) again all pointing to poor DNA recovery for these samples.

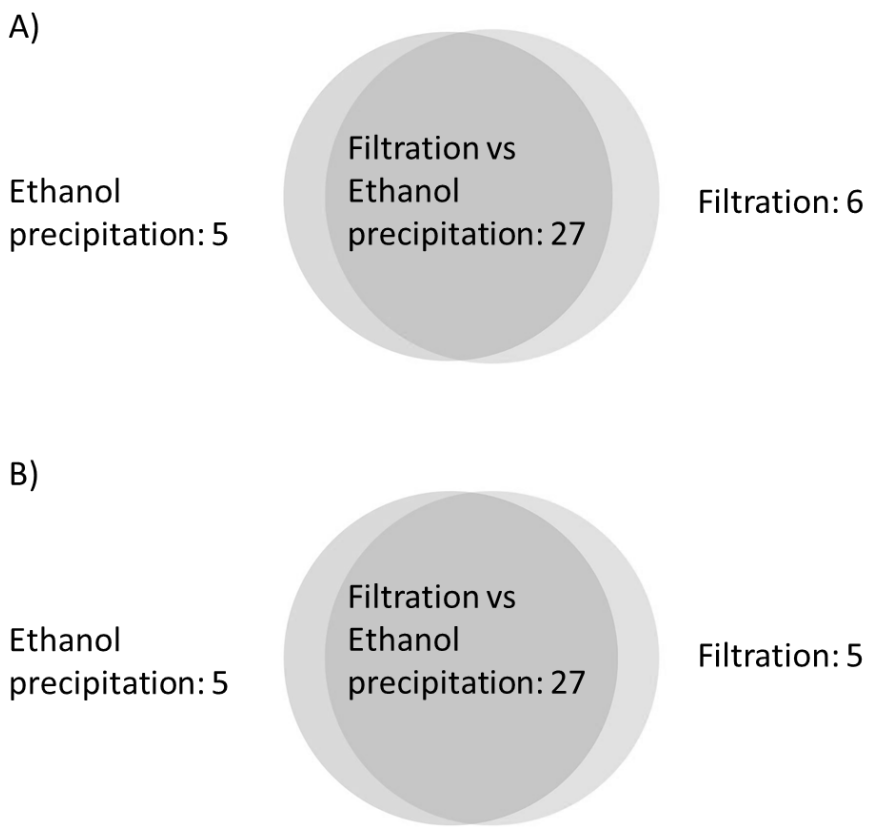

Figure 1. Venn diagram showing the number of GCN positive samples by one or both methods of sample collection. A) all samples included, B) missing ethanol precipitation results removed.

Metabarcoding PCR and library production

All 20 samples chosen for metabarcoding were successfully amplified, DNA quantified, and sent for indexing PCR and Illumina sequencing (Table 4). The 20 samples were selected to represent 14 ponds from across the HDBs that were positive for great crested newt with both precipitation and filtration methods, three ponds that only tested positive with the precipitation method and three ponds that only tested positive with the filtration method were also selected.

Bioinformatics and data analysis

Quality control reports generated by the sequencing service, Source Bioscience showed that both the indexed amplicons and the returned sequences were of good quality with all samples having a Q30 of above 70%. Most Illumina runs will generate >70-80% Q30 data and a high-quality score means that a base call is more reliable and less likely to be incorrect. Samples also had mean quality scores above 30 which is a measure of base calling accuracy equivalent to the probability of an incorrect base call 1 in 1000 times.

Sequencing results

For the 12S analysis, of the sequences that passed all filters and went onto taxonomic assignment a total of 1.32M sequences (~89% of sequences) were assigned a taxonomic identification which represented 29 species including fish, birds, mammals, and amphibians. A total of 161K sequences (~11% of sequences) were unassigned. The percentage of assigned and unassigned sequence reads per sample are shown in Figure 2. Results are shown in Table 5 and Appendix 2. Samples contained between two and 10 species of vertebrate with an average of 5 species per pond. The most found species were smooth newt, great crested newt and mallard duck. For the full list of 12S metabarcoding results, please see Appendix 2.

11 of the 20 GCN PCR positive samples chosen for metabarcoding were found to contain GCN via metabarcoding (Table 6). None of the samples that were only GCN positive via one of the sampling methods were found to be positive for GCN when using metabarcoding. Two of the samples which had high PCR scores via qPCR using both sampling methods (Z1-038 and B1-098) were also negative for GCN via metabarcoding which was unexpected given that of the 11 positive samples via metabarcoding all but one (C1-049) were from samples with high PCR scores. The remaining sample with low PCR scores via qPCR using both sampling methods (B1-121) was negative via metabarcoding.

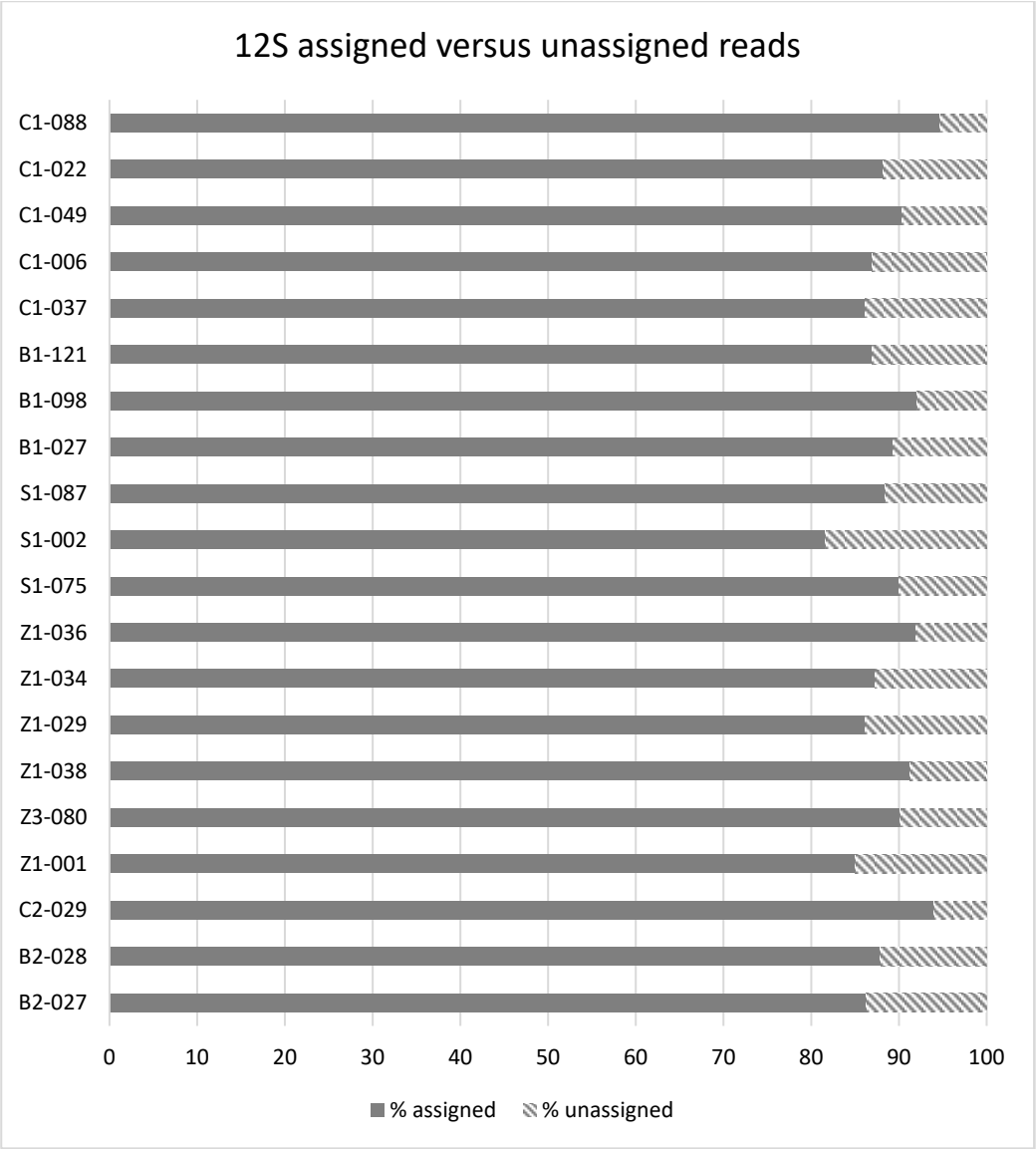

Figure 2. 12S proportion of assigned and unassigned sequence reads.

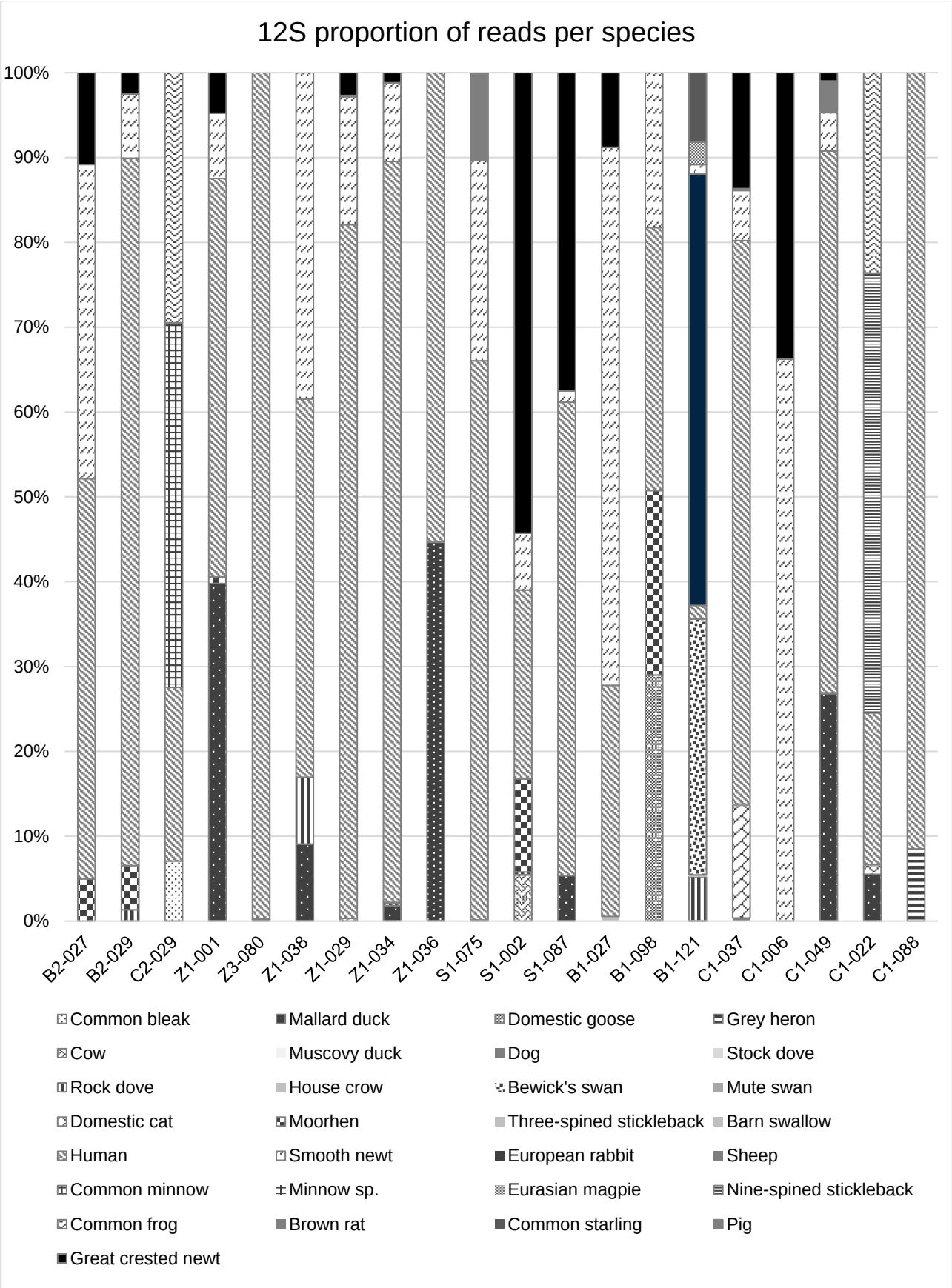

Figure 3. 12S proportion of reads per species

Table 5. Number of ponds containing species (vertebrates) using 12S primers.

Scientific name	Common name	no. of ponds
Alburns alburnus	Common bleak	1
Anas platyrhynchos	Mallard duck	8
Anser anser	Domestic goose	2
Ardea cinera	Grey heron	1
Bos taurus	Cow	2
Cairina moshata	Muscovy duck	4
Canis lupus familiaris	Dog	3
Columba oenas	Stock dove	1
Columba livia	Rock dove	7
Corvus splendens	House crow	1
Cygnus columbianus bewickii	Bewick's swan	1
Cygnus olor	Mute swan	1
Felis catus	Domestic cat	3
Gallinula chloropus	Moorhen	6
Gasterosteus aculeatus	Three-spined stickleback	2
Hirundo rustica	Barn swallow	1
Homo sapiens	Human	19
Lissotriton vulgaris	Smooth newt	15
Oryctolagus cuniculus	European rabbit	1
Ovis aries	Sheep	1
Phoxinus phoxinus	Common minnow	1
Phoxinus sp.		1
Pica pica	Eurasian magpie	1
Pungitius pungitius	Nine-spined stickleback	1
Rana temporaria	Common frog	2
Rattus norvegicus	Brown rat	1
Sturnus vulgaris	Common starling	1
Sus scrofa domesticus	Pig	4
Triturus cristatus	Great crested newt	11

Table 6. qPCR and metabarcoding results comparison

Sample name	GCN result filtration (number of positives out of 12)	GCN result ethanol precipitation (number of positives out of 12)	Metabarcoding result for GCN (numbers indicate read count where relevant)
FK1170, B2-027	12	9	Positive; 135566
FK1175, B2-029	6	11	Positive; 2503
FK1182, C2-029	0	1	Negative
FK1191, Z1-001	12	12	Positive; 1356
FK1192, Z3-080	0	1	Negative
FK1195, Z1-038	11	12	Negative
FK1196, Z1-029	11	10	Positive; 673
FK1205, Z1-034	12	12	Positive; 546
FK1212, Z1-036	3	0	Negative
FK1224, S1-075	3	0	Negative
FK1228, S1-002	12	9	Positive; 16817
FK1235, S1-087	12	12	Positive; 26228
FK1236, B1-027	12	2	Positive; 3936
FK1240, B1-098	12	8	Negative
FK1251, B1-121	2	5	Negative
FK1258, C1-037	12	12	Positive; 6437
FK1263, C1-006	11	12	Positive; 5430
FK1266, C1-049	5	2	Positive; 692
FK1268, C1-022	9	0	Negative
FK1269, C1-088	0	1	Negative

Note 1: Dark grey shading represents samples that are positive for GCN by qPCR using both sampling methods and metabarcoding.

Discussion

This work was undertaken to increase the evidence base for review of the current GCN eDNA monitoring protocols in terms of sample collection methodology. Despite 0.45 µM sterivex filters being used to collect the filtration samples, feedback from the habitat delivery bodies involved in sample collection raised concerns over the difficulty of pushing the pond water through the Sterivex filters especially those samples with suspended fine particles and/or algae with some samples not achieving the minimum level of 150 mL and

many not making it to 500 mL. It is known that pushing water through these filters due to turbidity issues can be difficult and this is one aspect discussed in previously published work considering the challenges of eDNA monitoring in freshwater ponds (Harper and others 2019). An alternative type of filter e.g. larger pore sized or involving a pre-filter could make sample collection easier. A larger pore sized filter/pre-filter is unlikely to become clogged as quickly with any sediment/particles present in the pond and eDNA loss for pore sizes up to 1 µm has been shown to be minimal (Turner and others. 2014, Wilcox and others. 2015).

In terms of single-species testing for GCN, the use of filtration and ethanol precipitation for sample collection were comparable with both giving approximately the same number of GCN positive results, especially when the additional positive results for filtration where no result was sent for ethanol precipitation were removed. For those four samples which were positive for GCN when collected using ethanol precipitation (C1-088, C2-029, S1-010, Z3-080), all four had 1 of 12 replicates being positive for GCN i.e. all four had a very low PCR score. One of these (C2-029) was achieved from only 78 mLs of filtered water (less than the 90 mLs collected when using ethanol precipitation) and also only 0.6 mLs of preservative was recovered which was on the low side when compared with the other samples which could explain the negative GCN result obtained for this sample when using filtration. Two of the four had 240 or 250 mLs water filtered respectively (C1-088, S1-010) which was less than the average volume filtered (325 mLs, all samples; 328 mLs, GCN positive samples), and the fourth (Z3-080) had 500 mLs water filtered. Of the six additional samples positive for GCN when collected via filtration (B1-108, B2-032 (2), C1-022, S1-075, Z1-036, Z1-023) there was a range of PCR scores (1, 1, 9, 3, 3 and 1 out of 12 respectively) and a range of water volumes filtered (350, 500, 500, 150, 500, and 300 mLs respectively). All six samples were from volumes of water of between 1.7 times and 5.6 times the volume collected via ethanol precipitation and the larger sample volumes could explain why these samples were positive for GCN when collected via filtration and negative when collected via ethanol precipitation.

For the metabarcoding of a sub-set of 20 samples chosen by Natural England, where possible this followed previous examples of similar metabarcoding work, instead of designing and trialling new PCR primers, which was beyond the scope of this project. Primers that had previously been described (and are in widespread use) were used to generate PCR amplicons from each sample (Riaz and others 2011). These primers were chosen as they are able to amplify a range of classes of Chordata including Amphibia, Aves, and Mammalia and have been successfully used in other Natural England studies (Rees and others 2023). In carrying out the first round PCR the aim was to capture as much of the sequence diversity as possible that is contained within the samples. Metabarcoding identified GCN in 11 of the 20 samples and a total of 29 vertebrate species across all samples analysed. It is thought that the eDNA score is a surrogate for the amount of target eDNA within a sample (Biggs and others 2014). Thus, the fact that all but one of these 11 positives were from samples with high PCR scores and likely containing high amounts of GCN eDNA suggests that metabarcoding is not sensitive enough to detect GCN in samples with the low levels of target eDNA that are associated with low

PCR scores. Metabarcoding has been shown to be less sensitive than qPCR for multiple species (Harper and others 2018, Hikaru and others 2018, Schenekar and others 2020, Yu and others 2022) and is a known drawback of metabarcoding. This explanation would also fit for the samples which were only GCN positive by one of the sampling methods or the samples with lower PCR scores not having GCN detected by metabarcoding as all except one (C1-022) had a PCR score of less than/equal to 5/12. Two of the samples which had high PCR scores via qPCR using both sampling methods (Z1-038 and B1-098) were negative for GCN via metabarcoding which was unexpected as they would be expected to contain larger amounts of GCN eDNA and thus be detectable via metabarcoding. These results suggest that without methodological advances such as the use of human blocking primers metabarcoding is not sensitive enough to be used to reliably detect GCN in pond eDNA samples.

The number of species found in some of the ponds was low with an average of five species per pond however, this was similar to the four species found on average (one to 17 species per pond) during the citizen science Genepools project (Rees and others 2023). All of the ponds were relatively new ponds ranging from one to four years of age, which could affect the number of species that you would expect to find, although new ponds do quickly become rich in species (Freshwater Habitats Trust). A study by Harper which used the same 12S primers as used here found 53 vertebrate species in total over 532 eDNA samples compared with our 29 species over 20 eDNA samples and 12 of these species were only found in one pond (Harper and others. 2020). Several species of bird were found as often the birds visit small ponds to drink and bathe. There are relatively few 12S sequences for birds within Genbank and this could, in part, be an explanation for some expected species not being found. The finding of Bewick's swan in one of the ponds (B1-121) would appear to be unusual given that this species tends to be a winter visitor in Eastern England and the Severn Estuary, however, the high read count (23070 or 30% of reads) suggests that this was not a sequencing error for example it is possible that this was from droppings that were stirred up out of the sediment in the sample collection. DNA from Red Jungle fowl was found, the 12S sequence for this species is very similar to that of domestic chicken so these records were amended in our results. Human DNA contamination of samples was an issue despite all samples being processed within a laminar flow cabinet equipped with UV decontamination. Human DNA accounted for ~50% of all assigned reads, and in the worst case accounted for 91% of the sequence reads in the sample. Human DNA could have been introduced at various parts along the process - from sampling pack assembly, contamination during sampling, and human DNA being present within the water body. To reduce this human DNA contamination, it is possible to use blocking primers (Seyama and others., 1992; Vestheim and Jarman, 2008; Craig and others., 2014) which effectively prevent human DNA from being amplified during the first round of the metabarcoding PCR. This should allow for more efficient amplification of the other species eDNA within the sample.

DNA from certain species can be misrepresented in the pool of eDNA - either DNA from species that are much smaller in size than others within the sample pool, or DNA from species present in much smaller numbers than the dominant species. Therefore, sequence read number does not necessarily correlate with species abundance. Additionally, DNA may

have been inefficiently extracted from different species and/or there was differential degradation of the DNA. The DNA of some individuals may also have degraded more than that of others due to different rates of degradation. DNA is liable to degradation by factors such as nucleases, UV light, microbial action and the temperature and humidity of storage conditions all of which will affect DNA quality after sample collection.

PCR amplification biases must also be considered in any metabarcoding study and these affect the ability of metabarcoding to give information on abundance of species. The primers that are used in the initial PCR may have missed some species due to biases and/or the primers used may simply not work efficiently for some species (Preston and others 2022). The success of metabarcoding is dependent upon the primer set chosen for use and its target loci and can be informed by in silico analysis. Ideally primers should target a hypervariable region (for high resolution taxonomic discrimination) and thus will determine the efficiency and accuracy of species detection and identification. DNA is in constant competition to bind to the primers during PCR amplification, and this competition can prevent the effective amplification of all species present as only the more common template DNAs are likely to be amplified (Kelly and others. 2014). This can mean that low abundance species are not detected termed ‘species masking’ (Brandon-Mong and others. 2015; Evans and others. 2016; Kelly and others. 2014). Metabarcoding may therefore be less capable of identifying the DNA of less abundant species within a community than a species-specific qPCR. It is possible that in this study, the high human content results in ‘species masking’ of GCN (and other species) within some of the samples.

The number of unassigned reads corresponded to ~11% of the total number of reads (ranging from 5.3 to 19.3% for individual samples). Unassigned reads are primarily due to a lack of available sequence data termed ‘gaps’ in the sequence databases which is a known problem (reviewed in Macadam and others. 2020) and it is inevitable that there will be unassigned reads in any metabarcoding study. Sequencing efforts such as the Darwin Tree of Life project which aims to generate DNA barcodes and full genomes for all UK species will help to alleviate this problem over the next several years and data generated can be reassessed as more sequence data becomes available.

Conclusions

The collection of pond water samples using filtration in general allowed a larger volume of water to be sampled, which resulted in the detection of GCN in six additional samples when compared with ethanol precipitation as a sample collection method, with most of these additional samples resulting in 1/12 PCR scores. However, it should be noted that ethanol precipitation also resulted in five samples where GCN was detected and these were not in the corresponding filtration samples. An alternative or larger pore sized filter could allow for easier sample collection as this was flagged by the HDB’s as a difficulty using the 0.45 µm Sterivex filters. The use of a larger pore size pre-filter could also assist in this matter.

As a method for the detection of GCN, metabarcoding was not as sensitive as qPCR which was as expected as in general only samples containing higher amounts of GCN eDNA were positive for GCN via metabarcoding. If human DNA reads were lower, it is possible that GCN may have been found in more samples as on average 50% of the read counts were human, in order to do this future work should look at using/developing human blocking primers for the 12S primers used in this study. It is also worth noting that this metabarcoding study was a carried out on a small scale which may require a larger scale repetition in the future. The metabarcoding study carried out by Harper and others 2018 which looked to determine if metabarcoding could be used to detect GCN in eDNA samples should also be taken into consideration as this analysed far larger numbers of samples and could help to inform a better understanding of the results.

References

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Griffiths, A.A., Foster, J., Wilkinson, J., Arnett, A., Williams, P., Dunn, F. (2014). Analytical and methodological development for improved surveillance of the Great Crested Newt (WC1067). Appendix 5. Technical advice note for field and laboratory sampling of great crested newt (*Triturus cristatus*) environmental DNA. Freshwater Habitats Trust, Oxford.

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

Brandon-Mong, G.J., Gan, H.M., Sing, K.W., Lee, P.S., Lim, P. E., Wilson, J.J. 2015. DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. Entomological Research, 105, 717-727. Doi:10.1017/S0007485315000681

Craig, C., Kimmerer, W. J., and Cohen, C. S. (2014). A DNA-based method for investigating feeding by copepod nauplii. J. Plankton. Res. 36, 271–275. doi: 10.1093/plankt/fbt104

Deiner, K., Lopez, J., Bourne, S., Holman, L., Seymour, M., Grey, E.K., Lacoursière, A., Li, Y., Renshaw, M.A., Pfrender, M.E., Rius, M., Bernatchez, L., Lodge, D.M. (2018). Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. Metabarcoding and Metagenomics 2. DOI: 10.3897/mbmg.2.28963

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

Ficetola, G.F., Miaud, C., Pompanon, F. Taberlet, P. (2008). Species detection using environmental DNA from water samples. Biology Letters, 4, 423-425. DOI: 10.1098/rsbl.2008.0118

Harper, L.R., Handley, L.L., Hahn, C., Boonham, N., Rees, H.C., Gough, K.C., Lewis, E., Adams, I.P., Brotherton, P., Phillips, S., Hanfling, B. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (Triturus cristatus). Ecology and Evolution, 8, 6330-6341. DOI: 10.1002/ece3.4013

Harper, L.R., Buxton, A.S., Rees, H.C., Bruce, K., Brys, R. Halfmaerten, D., Read, D.S., Watson, H.V., Sayer, C.D., Jones, E.P., Priestly, V., Machler, E., Murria, C., Garces-Pastor, S., Madupin, C., Urgess, K., Benson, G., Boonham, N., Griffiths, R.A., Handley, L.L., Hanfling, B. (2019). Prospects and challenges of environmental DNA (eDNA)

monitoring in freshwater ponds. Hydrobiologia, 826, 25–41. DOI: 10.1007/s10750-018-3750-5

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

Hosler, D. (2017). Where is the body? Dreissenid mussels, raw water testing, and the real value of environmental DNA. Management of Biological Invasions: International Journal of Applied Research on Biological Invasions 8, 335–341. DOI: 10.3391/mbi.2017.8.3.07

Illumina (2012). 16S Metagenomic Sequencing Library Preparation *Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System.*

Kelly, R.P., Port, J.A., Yamahara, K.M., Crowder, L.B. (2014). Using environmental DNA to census marine fishes in a large mesocosm. PLos ONE, 9, e86175.

MacAdam, C., Robins, J. & Thomson, T. (2020). 2020 Gap Analysis of the BOLD Database for Key English Invertebrates, Natural England Commissioned Report Number 32.

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. Freshwater Biology 63, 569-580. DOI: 10.1111/fwb.13094

Preston, M., Fritzsche, M., Woodcock, P. (2022). Understanding and mitigating errors and biases in metabarcoding: an introduction for non-specialists. JNCC Report No. 699, JNCC, Peterborough, ISSN 0963-8091.

Rees, H.C., Baker, C.B., Kane, S.D., Owen, J. 2023. GenePools – Investigating the life that dwells in garden ponds. Natural England Commissioned Reports, Number 454.

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

Sepulveda, A.J., Schabacker, J., Smith, S., Al-Chokhachy, R., Luikart, G., Amish, S.J. (2019). Improved detection of rare, endangered and invasive trout in using a new large-volume sampling method for eDNA capture. Environmental DNA 1: 227–237. DOI: 10.1002/edn3.23

Seyama, T., Ito, T., Hayashi, T., Mizuno, T., Nakamura, N., AND Akiyama, M. (1992). A novel blocker-PCR method for detection of rare mutant alleles in the presence of an excess amount of normal DNA. Nucleic Acids Res. 20, 2493–2496. doi: 10.1093/nar/20.10.2493

Schenekar, T., Schletterer, M., Lecaudey, L.A., Weiss, S.J. (2020). Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: Lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. River Research and Applications, 36, 1004-1013. DOI: 10.1002/rra.3610

Thalinger, B., Deiner, K., Harper. L. R., Rees, H.C., Blackman, R.C., Sint, D., Traugott, M., Goldberg, C., Bruce, K. 2021. A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. Environmental DNA, 3, 823-836. DOI: 10.1002/edn3.189

Thomsen, P.T., Kielgast, J., Iverson, L.L., Wius, C., Rasmussen, M., Gilbert, M.T.P., Orlando, L., Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology, 21, 2565-2573. DOI: 10.1111/j.1365-294X.2011.05418.x

Turner, C.R., Barnes, M.A., Xu, C.C.Y., Jones, S.E., Jerde, C.L., Lodge, D.M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 2014: 676–684. DOI: 10.1111/2041-210X.12206

Vestheim, H., and Jarman, S. N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. Front. Zool. 5:12. doi: 10.1186/1742-9994-5-12

Wilcox, T.M., McKelvey, K.S., Young, M.K., Lowe, W.H., Schwartz, M.K. (2015). Environmental DNA particle size distribution from Brook Trout (Salvelinus fontinalis). Conservation Genetics Resources 7: 639–641. DOI: 10.1007/s12686-015-0465-z

Yu, Z., Ito, S.I., Wong, M.K., Yoshizawa, S., Inoue, J., Itoh, S., Yukami, R., Ishikawa, K., Guo, C., Ijichi, M., Hyodo, S. (2022). Comparison of species-specific qPCR and metabarcoding methods to detect small pelagic fish distribution from open ocean environmental DNA. PLoS One, 17(9):e0273670. DOI: 10.1371/journal.pone.0273670.

List of tables

List of figures

Appendix 1

DNA extraction from sterivex filters

All surfaces were cleaned with bleach solution prior to commencing DNA extraction and then periodically during DNA extraction process.

1. ATL and AL Buffers were pre-warmed at 56°C.
2. An appropriate amount of ATL and AL buffers were pre-mixed with each filter requiring: 540 µL ATL and 300 µL AL. Additionally, a piece of control DNA was added at a known concentration to allow monitoring of DNA extraction efficiency.
3. An individual filter sample was removed from its container and the outside wiped down with bleach solution.
4. Preservative was removed from the filter (after removal of the inlet and outlet caps) into 1. 5 mL microcentrifuge tube/s using a fresh sterile 10 mL luer lock syringe for every filter. The volume recovered was recorded and retained for future use.
5. The outlet cap was replaced and 840 µL of pre-warmed ATL/AL solution and 50 µL proteinase K was added to the filter before replacing the inlet cap.
6. Steps 3 to 5 were repeated on each filter sample, changing gloves between each filter sample.
7. An extraction blank was set up on each DNA extraction day by adding 840 µL of the pre-warmed ATL/AL solution and 50 µL Proteinase-K to a clean 1.5 mL microcentrifuge tube.
8. All filters (and extraction blank) were briefly vortexed before being placed into 50 mL centrifuge tubes and incubated at 56°C in a water bath for 1 hour.
9. All filters were briefly vortexed every 10 minutes to ensure even and thorough digestion of material on all parts of the filter.
10. Whilst incubating filters, the microcentrifuge tube tubes containing the expelled preservative were centrifuged at 13,000 x*g* for 30 min at room temperature to pellet any DNA or material present.
11. The supernatant was removed using a pipette (avoiding pellet) and tubes retained for step 12.
12. Using the appropriate luer lock syringes from step 4 the digestion mixture was expelled from the filter (after removal of inlet and outlet caps) into the microcentrifuge tube with corresponding preservative pellet before briefly vortexing to resuspend the pelleted material.
13. The outlet cap was replaced and 500 µL 100% ethanol to each filter before replacing the inlet cap and briefly vortexing the filters.
14. The ethanol was expelled into the corresponding microcentrifuge tube containing the digest solution using the appropriate luer lock syringes from step 4/12 (after removal of inlet and outlet caps) and the microcentrifuge tube briefly vortexed to thoroughly mix.
15. The digest mixture was added onto a DNeasy spin column in 650 µL volumes (repeated until the entire extract has been passed through the spin column) and centrifuged at 6,000 xg for 1 minute.
16. The spin column was placed into a new 2 mL collection tube and the flow-through discarded.

17. 500 µL buffer AW1 was added and centrifuged for 1 minutes at 6,000 xg.
18. The spin column was placed into a new 2 mL collection tube and the flow-through discarded.
19. 500 µL buffer AW2 was added and centrifuged for 3 minutes at 20,000 xg. The flow-through was discarded and spin columns were re-centrifuged for 1 minute to dry the column membrane.
20. The spin columns were transferred to pre-labelled 1.5 m microcentrifuge tubes.
21. DNA was eluted by the addition of 200 µL AE buffer before incubating at room temperature for 1 minute and centrifugation for 1 minute at 6,000 xg.
22. The DNA samples were aliquoted in 4 equal amounts before storage at -20 °C.

DNA Quantification

DNA extracts were quantified using the Qubit® dsDNA BR assay kit and Qubit 3.0 fluorimeter as follows:

1. The Qubit® working solution was prepared by diluting the Qubit® dsDNA BR reagent 1:200 in Qubit® dsDNA BR buffer.
2. Make up two standards by adding 190 µL Qubit® working solution into each of two tubes before adding 10 µL of each Qubit® standard to the appropriate tube. Mix by vortexing.

For each extract make up a tube with a final volume of 200 µL containing 1-20 µL extract and 180-199 µL Qubit® working solution

DNA extraction efficiency PCR

PCRs were set up in a total volume of 25 µL consisting of:

a. 3 µL of extracted template DNA,
b. 1 µL of each primer/probe (0.4 µmol/L DegL; 0.4 µmol/L DegR; 0.1 µmol/L Deg.probe),
c. 12.5 µL of TaqMan® Environmental Master Mix 2.0 (containing AmpliTaq GOLD DNA polymerase),
d. 6.5 µL ddH2O.

Each sample was run in duplicate and each plate included 8 positive controls (4 replicates each at $1x10^{-3}$ and $1x10^{-4}$ ng/µl synthetic control DNA) and 4 negative controls (ultrapure water in place of DNA) on a Bio-Rad CFX Connect real-time PCR machine as follows: an initial incubation for 5 minutes at 56.3ºC then 10 minutes at 95°C; followed by 35 cycles with a melting temperature of 95°C for 30 seconds and an annealing temperature of 52ºC for 1 minute.

Species-specific GCN qPCR

PCRs were set up in a total volume of 25 µL consisting of:

e. 3 µL of extracted template DNA at 1 ng/µL,

f. 1 µL of each primer/probe (0.4 µmol/L TCCBL; 0.4 µmol/L TCCBR; 0.1 µmol/L TCCB.probe),
g. 12.5 µL of TaqMan® Environmental Master Mix 2.0 (containing AmpliTaq GOLD DNA polymerase),
h. 6.5 µL ddH2O.

Each sample was run as 12 replicates and each plate included 8 positive controls (4 replicates each at $1x10^{-3}$ and $1x10^{-4}$ ng/µl GCN DNA) and 4 negative controls (ultrapure water in place of DNA) on a Bio-Rad CFX Connect real-time PCR machine as follows: an initial incubation for 5 minutes at 56.3ºC then 10 minutes at 95°C; followed by 55 cycles with a melting temperature of 95°C for 30 seconds and an annealing temperature of 52ºC for 1 minute.

Nucleospin® gel and PCR cleanup

For DNA extraction from agarose gels:

1. Excise DNA fragment from gel with a fresh sterile scalpel blade for each sample.
2. Determine the weight of the gel slice and add 200µl buffer NTI for every 100mg of agarose gel
3. Incubate for 5-10 minutes at 50°C vortexing every 2-3 minutes until the gel slice is completely dissolved.
4. Place a NucleoSpin® Gel and PCR clean-up column into a collection tube and load 700µl of sample onto the spin column and centrifuge for 30 seconds at 11,000 xg.
5. Wash the silica membrane by adding 700 µL Buffer NT3 to the column and centrifuge for 30 seconds at 11,000 xg.
6. Discard the flow-through and place the column back into the collection tube before repeating this wash step.
7. Dry the silica membrane for one minute at 11,000 xg to remove Buffer NT3 completely.

Elute the DNA by placing the column into a fresh 1.5 mL microcentrifuge tube and add 20 µL Buffer NE and incubate at room temperature for one minute before centrifuging for one minute at 11,000 xg

Index PCR

PCRs were set up in a total volume of 50 µL consisting of:

a. 25 µl 2x KAPA HotStart ReadyMix
b. 5 µl Nextera XT Index 1 Primers
c. 5 µL Nextera XT Index 2 Primers
d. 10 µL PCR grade water
e. 5 µL DNA

PCR cycling was as follows: an initial incubation for 3 minutes at 95°C; followed by 8 cycles with denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 30 seconds; and a final extension step at 72°C for 5 minutes before holding at 4°C until collection of PCR products for analysis.

Appendix 2. 12S metabarcoding results

Note that many cells have been left blank intentionally.

Scientific name	Common name	FK170 B2-027	FK175 B2-029	FK182 C2-029	FK191 Z1-001	FK192 Z3-080	FK195 Z1-038	FK196 Z1-029	FK205 Z1-034	FK212 Z1-036	FK224 S1-075	FK228 S1-002	FK235 S1-087	FK236 B1-027	FK240 B1-098	FK251 B1-121	FK258 C1-037	FK263 C1-006	FK266 C1-049	FK268 C1-022	FK269 C1-088
Alburns alburnus	Common bleak			10254																	
Anas platyrhynchos	Mallard duck				11301	26	12969		824	34275			3738						19205	707	
Anser anser	Domestic goose														20469	21					
Ardea cinera	Grey heron																				6923
Bos taurus	Cow											1676							48		
Cairina moshata	Muscovy duck				6					37			5						15		
Canis lupus familiaris	Dog								13			20					168				
Columba oenas	Stock dove															6					
Columba livia	Rock dove		1318			9	11253		160		113	99				3964					
Corvus splendens	House crow															174					
Cygnus columbianus bewickii	Bewick's swan															23070					

Scientific name	Common name	FK170 B2-027	FK175 B2-029	FK182 C2-029	FK191 Z1-001	FK192 Z3-080	FK195 Z1-038	FK196 Z1-029	FK205 Z1-034	FK212 Z1-036	FK224 S1-075	FK228 S1-002	FK235 S1-087	FK236 B1-027	FK240 B1-098	FK251 B1-121	FK258 C1-037	FK263 C1-006	FK266 C1-049	FK268 C1-022	FK269 C1-088
Cygnus olor	Mute swan																		5		
Felis catus	Domestic cat							78									6259			134	
Gallinula chloropus	Moorhen	6727	5181		220				17			3396			15346						
Gasterosteus aculeatus	Three-spined stickleback																70			16	
Hirundo rustica	Barn swallow													240							
Homo sapiens	Human	63961	82740	29646	13343	12684	63644	20546	40152	42464	58000	6903	39019	12218	21865	1236	31293		45767	2288	74178
Lissotriton vulgaris	Smooth newt	50208	7483		2187		54928	3789	4267		20832	2079	931	28434	12890	38909	2785	10648	3216		
Oryctolagus cuniculus	European rabbit															849					
Ovis aries	Sheep						9														
Phoxinus phoxinus	Common minnow			62215																	
Phoxinus sp.				41																	
Pica pica	Eurasian magpie															2027					
Pungitius pungitius	Nine-spined																			6640	

Scientific name	Common name	FK170 B2-027	FK175 B2-029	FK182 C2-029	FK191 Z1-001	FK192 Z3-080	FK195 Z1-038	FK196 Z1-029	FK205 Z1-034	FK212 Z1-036	FK224 S1-075	FK228 S1-002	FK235 S1-087	FK236 B1-027	FK240 B1-098	FK251 B1-121	FK258 C1-037	FK263 C1-006	FK266 C1-049	FK268 C1-022	FK269 C1-088
	stickle back																				
Rana temporaria	Common frog			42667																3026	
Rattus norvegicus	Brown rat																		113		
Sturnus vulgaris	Common starling															6264					
Sus scrofa domesticus	Pig							38			9066						102		2579		
Triturus cristatus	Great crested newt	14670	2503		1356			673	546			16817	26228	3936			6437	5430	692		
	Total read counts	135566	99225	144823	28413	12719	142803	25124	45979	76776	88011	30990	69921	44828	70570	76520	47114	16078	71640	12811	81101

www.gov.uk/natural-england