UNIVERSITY OF
Hull International
Fisheries Institute
HIFI

# Oxford FAS fisheries survey, October 2016

**FINAL REPORT TO ENVIRONMENT AGENCY**

**N. Angelopoulos, R. A. A. Noble, A. D. Nunn, J. P. Harvey, J. D. Bolland, H. V. Watson, J. Li, B. Haenfling & I. G. Cowx**

**Hull International Fisheries Institute,**
**University of Hull,**
**Hull, HU6 7RX**

**March 2017**

UNIVERSITY OF Hull

**TABLE OF CONTENTS** **Page No.**

## LIST OF TABLES



## LIST OF FIGURES

**LIST OF PLATES** **Page No.**



**LIST OF APPENDICES** **Page No.**

**ACKNOWLEDGEMENTS**
The authors would like to thank Stuart Manwaring and Edith Putt for organising access to the site, and Jamie Dodd and Leona Murphy for assistance in the field. The study was funded by the Environment Agency.

## EXECUTIVE SUMMARY

The Oxford Flood Alleviation Scheme (FAS) proposes a new channel to be excavated to the west of Oxford to reduce the risk of flooding to property and critical infrastructure. A number of environmental surveys will be undertaken within the study area to provide the project team with information to help understand the baseline environment. Fisheries surveys were undertaken to inform the Oxford FAS assessments, to understand the fish populations associated with the watercourses to the west of Oxford, and to identify invasive non-native species that might be present and have the potential to spread in to the new excavated flood channel. These surveys form a key part of the Water Framework Directive assessment (to be undertaken by others) and allow informed decisions to be made about the treatment of each watercourse that may be affected by the scheme.

Sites 1 to 6 of Oxford FAS (Seacourt stream, Botley stream, Bulstake stream, Hinksey stream) appear to support fish assemblages expected for lowland streams. The assemblages are dominated by eurytopic ('generalist') species, particularly roach, chub, perch and bleak and this was probably due to the availability of suitable habitat such as slow to moderate flows, areas of deep water, bankside vegetation and silt/sand or gravel substrates. Roach was consistently in the top 2 most abundant species across sites 1 to 6; other species that accounted for <10 % of fish captured at a number of sites were common bream, dace, gudgeon, minnow, ruffe, bullhead, tench and roach/bream hybrid. The low number of juvenile fish recorded for all species at all sites could indicate potentially poor recruitment, but it is possible that the densities of juvenile fish were underestimated due to low sampling efficiency by electric fishing because the water was deep and in some areas the river was wide and marginal macrophyte growth was dense. Site 7 Chilswell drain was extremely overgrown which limited the river length that could be sampled to 20 m and no fish were captured. Chilswell drain will be more suitable to sample in the spring before the bankside vegetation has grown.

Across all samples eDNA of ten species was detected: Common Bream, Silver Bream (*Blicca bjoerkna*), Common Carp (*Cyprinus carpio*), Pike, Ide (*Leuciscus idus*), Perch, Roach, Rudd (*Scardinius erythrophthalmus*), Chub and Tench. The six most common species were detected with both markers (highest confidence of presence: *A. brama*, *B. bjoerkna*, *E. lucius*, *P. fluviatilis*, *R. rutilus*, and *T. tinca*), whereas three species were only detected above the filtering threshold with one of the two markers (CytB: *C. carpio* and *S. erythrophthalmus*; 12S: *S. cephalus*). One species was present at very low read counts with only one marker (12S: *L. idus*). When applying the filtering criteria, *L. idus* was no longer present and *S. cephalus* was absent for CytB. The highest total number of read counts was recorded for *E. lucius*, *P. fluviatilis* and *R. rutilus* with a total of over 100,000 reads for at least one of the two markers.

# 1. INTRODUCTION

The Oxford Flood Alleviation Scheme (FAS) proposes a new channel to be excavated to the west of Oxford to reduce the risk of flooding to property and critical infrastructure. A number of environmental surveys will be undertaken within the study area to provide the project team with information to understand the baseline environment. Surveys are required to assess the ecological status in the affected watercourses, and will form a key part of the Water Framework Directive assessment (to be undertaken by others) and allow informed decisions to be made about the treatment of each watercourse that may be affected by the scheme.

Baseline fisheries surveys and eDNA survey are therefore required to inform the Oxford FAS assessments, to help understand the fish populations associated with the watercourses to the west of Oxford and to identify invasive non-native species that might be present and have the potential to spread in to the new excavated flood channel. Data will be gathered on fish species and abundance by carrying out the following surveys:

1. Determining fish density at seven sites by three run catch-depletion surveys using electric fishing
2. Identifying presence and relative abundance of fish communities in Kennington Pond using environmental DNA (eDNA) based metabarcoding

# 2. MATERIALS AND METHODS

## *2.1 Fisheries methodology*

Fisheries surveys were carried out at the study sites listed in Table 1 (Figure 1) in late October 2016, using electric fishing conducted in accordance with the EA's Operational Instructions "Electric Fishing Operations: Equipment and Working Practices" using a 2 kVA generator with an Easyfisher variable-output control box. Where possible the sites were 100 m long or ten times as long as the river width, whichever was the greatest. Quantitative electric fishing surveys (estimates of absolute abundance based on three-catch removal method; see Zippin (1956) and Carle & Strub (1978)) were conducted by three operatives (one anode operator and two netsmen) fishing in an upstream direction. The fourth operator supervised safe operation of the electric fishing equipment that was pushed along in a small boat. Each survey reach was isolated using upstream and downstream stop nets, to ensure no escape from, or migration into, the sampling area. Immobilised fishes were captured using hand-nets and transferred to water-filled, aerated containers prior to data collection. All fish were identified to species level, measured (fork-length, nearest mm) and scales taken (for possible future ageing), before being returned to the river.

At Site 7 it was only possible to sample 20 m of river in a single run with no stop nets due to overgrown vegetation restricting wider access to a reach of 100 m length.

Density estimates for each species at each site were derived from estimates of absolute abundance based on the three-catch removal method and estimates of populations were calculated by the Maximum Likelihood Method (Carle & Strub 1978). In all cases the population densities were expressed as numbers/100 $m^2$. When catches were sufficient, length-frequency distributions were derived to examine the size structure of the populations. Fish scales were not aged as it was out of the scope of this project.

**Table 1.** Fisheries survey sites sampled October 2016

| Site | Site name | Date | NGR | Length (m)/Average width (m)/area ($m^2$) |
|---|---|---|---|---|
| 1 | Seacourt stream | 25/10/2016 | SP4915706585 | 100/6/600 |
| 2 | Seacourt stream park & ride | 25/10/2016 | SP4916006583 | 60/5/300 |
| 3 | Botley stream by Golf Range | 27/10/2016 | SP4950006522 | 100/5/500 |
| 4 | Seacourt stream at North Hinksey | 27/10/2016 | SP4923105784 | 100/5/500 |
| 5 | Bulstake Stream | 26/10/2016 | SP4985605727 | 100/8/800 |
| 6 | Hinksey stream | 26/10/2016 | SP5049205102 | 100/5/500 |
| 7 | Chilswell drain | 26/10/2016 | SP5118404103 | 20/2/40* |

**single run*

### *2.2 eDNA Method*

#### *2.2.1 Water sample collection*

Water samples were collected from Kennington Pond 28th October 2016 (Figure 1). Collection and filtration of samples were carried out according to the approach developed in our EA and SEPA (Scottish Environment Protection Agency) funded projects "Development of an eDNA metabarcoding assay for Water Framework Directive Phase 1 and 2" (Hänfling et al. 2016). Samples were collected by hand from the shoreline and consisted of 5 x 400 ml subsamples filling a sterile 2 litre (L) collection bottle, taken within a range of 50 m. Due to the size (0.45 hectares) and spatial complexity (numerous semi-isolated bays and subsections) of Kennington Pond, 10 x 2 L samples were collected. Two blanks (purified water) were included at the filtration stage. All samples were stored chilled and filtered on the day of collection (N = 12).

#### *2.2.2 Water filtration*

In a dedicated eDNA facility at the University of Hull, 500 ml of each water sample was filtered through sterile 0.45 µm cellulose nitrate membrane filters and pads (47 mm diameter; Whatman, GE Healthcare, UK) using Nalgene filtration units in combination with a vacuum. A second filter was required for all Kennington Pond samples, as these were slow to filter. Only one filter was required for each blank. All filters were stored at -20 °C until DNA extraction.

#### *2.2.3 DNA extraction and PCR*

DNA was extracted from the filters using the PowerWater DNA Isolation Kit (MO BIO Laboratories, Inc. Carlsbad, USA. The kit is now sold as: DNeasy PowerWater Kit, Qiagen, DE) following the manufacturer's instructions. DNA samples were quantified via fluorometer using the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, US).

Sequencing libraries were generated from PCR amplicons of loci from two mitochondrial regions: 12S and Cytochrome B (CytB). To enable the detection of possible PCR contamination, no-template controls (NTCs) and single-template controls (STCs), in which the PCR reaction is set-up using molecular grade water or cichlid fish DNA respectively, were included as additional samples within each library.

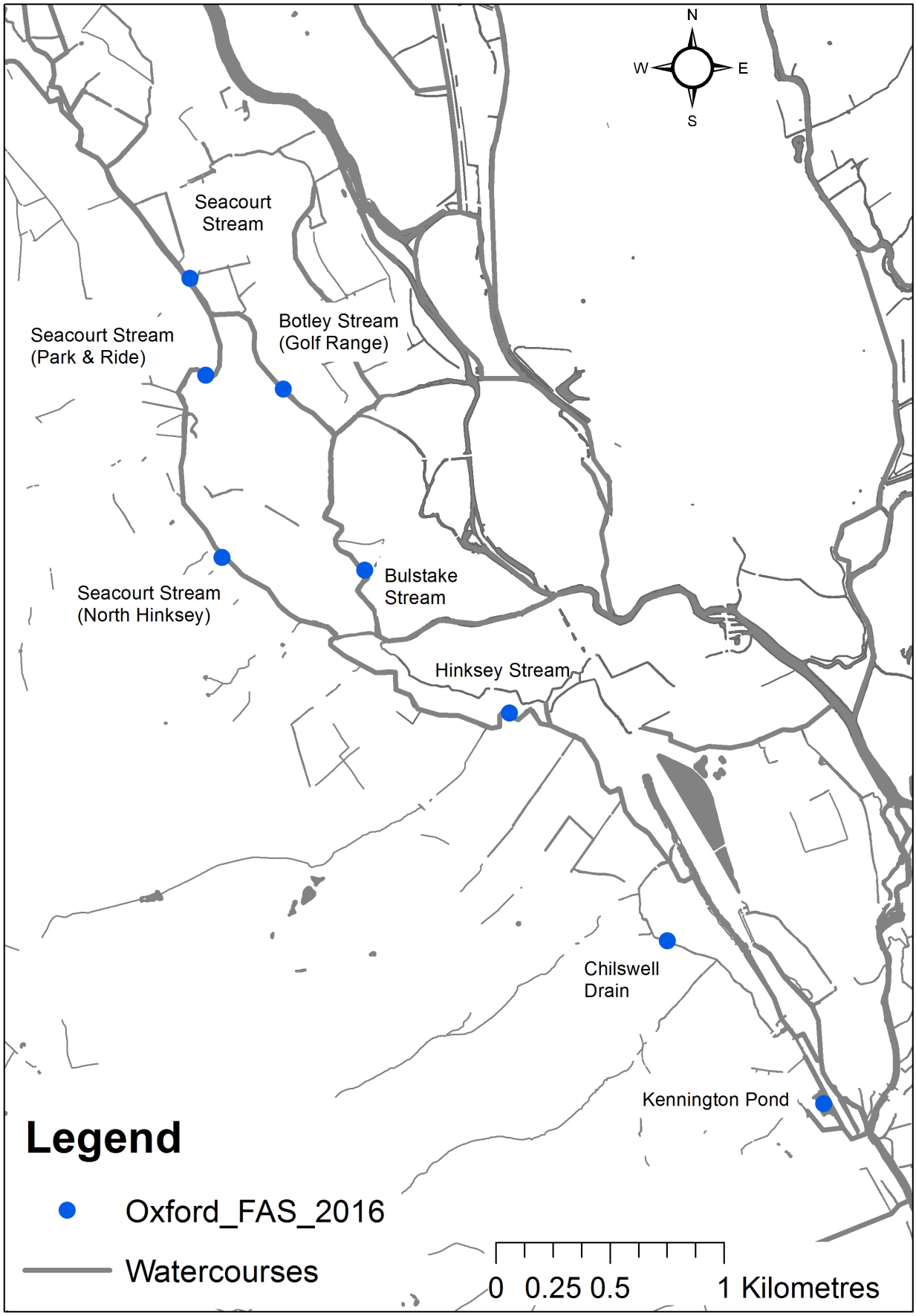


**Figure 1.** Fisheries survey and eDNA locations for the Oxford FAS.

All of the samples (N = 12) were performed in three replicates to allow detection thresholds to be determined. This enabled us to assign a level of confidence to the species detected (for example, if a species was detected in 3/3 replicates, we can have full confidence that it is present). These three replicates were pooled based on the loci (12S/CytB) and were sequenced on an Illumina MiSeq for 'FastQC' generation only following the Illumina guidelines for MiSeq.

The PCR methodology varied between the loci/library and is described here per locus:

*2.2.3.1 CytB*

The CytB locus was amplified using the 'one-step' protocol employed in our previous eDNA fish-metabarcoding studies (Hänfling et al. 2016). Using already developed primers (Kocher et al. 1989), this single PCR reaction targets a 460 base pair (bp) region of the CytB gene, and simultaneously includes the adapters required for DNA sequencing and an index tag, enabling the identification of sequences generated from individual samples in the resulting data.

The 20 µl volume PCR reaction included: 10 µl Q5® Hot Start High-Fidelity 2X Master Mix (New England Biolabs, UK), 1 µl forward primer (10 µM), 1 µl reverse primer (10 µM), 2 µl DNA, and 6 µl molecular grade water. A 'step-down' cycling protocol was incorporated to allow for potential mismatches across a range of taxa. Thermal cycling parameters were as follows: (i) 98 °C for 2 min; (ii) 98 °C for 10 s; (iii) 55°C for 20 s; (iv) 72 °C for 30 s; (v) repeat steps 2-4 an additional nine times; (vi) 98 °C for 10 s; (vii) 53°C for 20 s; (viii) 72 °C for 30 s; (ix) repeat steps 6-8 an additional nine times; (x) 98 °C for 10 s; (xi) 50°C for 20 s; (xii) 72 °C for 30 s; (xiii) repeat steps 10-12 an additional 29 times; (xiv) 72 °C for 10 min; (xv) hold at 4 °C. The successful amplification of specific PCR amplicons was confirmed by visualisation of PCR products via gel electrophoresis. Each PCR reaction was performed in triplicate repeat, which were subsequently pooled prior to the library preparation steps described below.

*2.2.3.2 12S*

Following the consistently lower sequencing yield of the one-step protocol for the 12S locus compared to CytB, it was decided to switch to a previously successful two-step protocol, which involves two successive PCR reactions.

In the two-step protocol, the first PCR reaction setup was the same as that for the one-step 12S PCR, but used different 12S specific primers which do not contain sequencing adapters or index tags. A 'step-down' cycling protocol was incorporated to allow for potential mismatches across a range of taxa. Thermal cycling parameters were as follows: (i) 98 °C for 2 min; (ii) 98 °C for 10 s; (iii) 63°C for 20 s; (iv) 72 °C for 30 s; (v) repeat steps 2-4 an additional nine times; (vi) 98 °C for 10 s; (vii) 60°C for 20 s; (viii) 72 °C for 30 s; (ix) repeat steps 6–8 an additional nine times; (x) 98 °C for 10 s; (xi) 58°C for 20 s; (xii) 72 °C for 30 s; (xiii) repeat steps 10–12 an additional 20 times; (xiv) 72 °C for 10 min; (xv) hold at 4 °C. Each PCR reaction was performed in triplicate repeat.

The triplicate PCR products of each sample were then pooled and subsequently cleaned using the Mag-Bind® RXNPure Plus Kit (Omega Bio-tek, Inc. US) according to the manufacturer's guidelines.

The second PCR reaction attaches the adapter sequences required for sequencing and the index tags to the first PCR amplicon. The reaction setup was the same as the first PCR, and used primers specific to the first PCR amplicon which contained index tags. Thermal cycling parameters were as follows: initial denaturation at 95 °C for 3 min, followed by 12 cycles of 98 °C for 20 s and 72 °C 30 s, with a final extension of 72 °C for 5 min.

*2.2.4 Library preparation and sequencing*

Following the generation of PCR amplicons with sequencing adapters and a unique combination of dual-indexes per sample, each library was normalised for concentration across the samples using the SequalPrep Normalization Plate Kit (Invitrogen, Life Technologies) and the samples then pooled. Each library was size separated via gel electrophoresis and subsequently extracted from the gel, allowing for the removal of any non-specific PCR products. Libraries were then quantified by qPCR, using the NEB-Next Library Quant Kit (New England Biolabs, UK) and diluted to a working concentration of 4 nM for sequencing. Both libraries were sequenced on an Illumina MiSeq for 'FastQC' generation only following the Illumina guidelines for MiSeq.

*2.2.5 Data Analysis*

Sequence data was analysed with our previously developed bioinformatics pipeline (see *Bioinformatics* below). This pipeline included a comprehensive fish reference database, which allowed the assignment of all sequence reads to individual species. In order to exclude the possibility of false positives, a two-step approach was used. Records for a taxon in a specific sample were only regarded as "true" if they (a) exceed a certain threshold value (proportion of reads in the sample) i.e. a proportion higher than 0.1% and 0.2% of all sequence reads in the sample for 12S and CytB respectively - these thresholds were experimentally determined in a previous study (Hänfling et al. 2016) to reduce the occurrence of false positives in control samples by over 90% - and (b) are detected in at least two of the three sequence replicates of that sample. Data on all true records are summarised in two ways, which serve as proxies of relative abundance: (a) Site Occupancy (SO, proportion of sample from where the taxon was confirmed) and (b) Read Counts (RC, total number of taxon specific sequences across all samples).

*2.2.6 Bioinformatics*

The bioinformatics analysis was carried out using the pipeline developed in Hänfling et al. (2016) but included a number of improvements. The approach relies on a comprehensive reference database for European freshwater fish and involves the following steps. The program Trimmomatic 0.32 (Bolger et al. 2014) was used for quality trimming and removal of adapter sequences from the raw Illumina reads. Average read quality was assessed in sliding windows (window size 5 bp) starting from the 3'-end of the read and reads were clipped until the average quality per window was above a phred base quality of 30 (equivalent to 99.99% accuracy). Subsequently, all reads shorter than a defined minimum read length (12S - 90 bp; CytB - 100 bp) were discarded. Sequence pairs were then merged into single high quality reads using the program FLASH 1.2.11 (Magoč & Salzberg 2011). To remove redundancy, sequences were clustered at 100% identity using vsearch 1.1 (https://github.com/torognes/vsearch). Any singletons, i.e. sequences occurring in only a single copy, were considered sequencing errors and omitted from further analyses. Non-redundant sets of query sequences were then compared to the respective curated non-redundant reference database using the Basic Local Alignment Search Tool, BLAST (Zhang et al. 2000). BLAST output was analysed using a custom python script, which implements a lowest common ancestor (LCA) approach for taxonomic assignment similar to the strategy used by the programme MEGAN (Huson et al. 2007). In brief, after the BLAST search the most significant matches were recorded to the reference database (yielding the top 10% bit-scores) for each of the query sequences. If only a single taxon was present in the top 10% the query was assigned directly to this taxon. If more than one reference taxon was present in the top 10%, the query was assigned to the lowest taxonomic level that was shared by all taxa in the list of most significant hits for this query. Sequences for which the best BLAST hit had a bit score below 80 or had less than 100%/95% identity (12S/CytB)

to any sequence in the curated database were considered non-target sequences. These were subjected to a separate BLAST search against the complete nucleotide database on Genbank.

## 3. RESULTS

### *3.1 Fisheries results*

A total of 13 fish species were captured across all seven sites; ≥ 9 species of fish were present at sites 1, 2, 3 and 5 but ≤ 6 species at sites 4 and 6 (Table 1). No fish were captured at site 7. Roach (*Rutilus rutilus* (L.)), chub (*Leuciscus cephalus* (L.), pike (*Esox lucius* L.) and perch (*Perca fluviatilis* L.) were present at all sites but abundance within each site varied. Bleak (*Alburnus alburnus* (L.)) densities were highest at site 1 (42 fish/100 m²) compared with sites 2 and 3 (3 fish/100 m²), 4 (1.2 fish/100 m²) and 5 (0.1 fish/100 m²); no bleak were captured at site 6 (Table 2, Figure 2b). Roach densities were highest at sites 2 (17.7 fish/100 m²), 1 (15.5 fish/100 m²) and 4 (15.4 fish/100 m²) and were ≤ 10 fish/100m² at sites 3, 5 and 6 (Table 2, Figure 2b). Chub, pike and perch were present at all sites with densities ≤ 10 fish/100m² while minnow (*Phoxinus Phoxinus* L.)), gudgeon (*Gobio gobio* (L.), dace (*Leuciscus leuciscus* (L.)), common bream (*Abramis brama* (L.)), bullhead (*Cottus gobio* L.), ruffe (*Gymnocephalus cernuus* (L.)) and tench (*Tinca tinca* (L.)) densities were ≤ 5 fish/100m² and these species were not present at all sites (Table 2, Figure 2b).

Bleak was the most abundant species at site 1, accounting for 59 % of fish captured, followed by roach (24 %) (Figure 2a). At site 2 roach (50 %) was the most abundant species followed by chub (19 %) and perch (11 %) (Figure 2a). Roach and chub were both equally abundant at site 3, each accounting for 25 % of fish captured, followed by perch (11 %) (Figure 2a). Roach was the most abundant species at site 4, accounting for 80 % of fish captured, whilst chub (42 %) and roach (34 %) were the most abundant species at site 5 (Figure 2a). Perch was the most abundant species at site 6, accounting for 35 % of fish captured, followed by roach (30 %), pike (20 %) and chub (15 %) (Figure 2a). Other species accounted for <10 % of fish captured at a number of sites and included common bream, dace, gudgeon, minnow, ruffe, bullhead, tench and roach/bream hybrid (Figure 2a).

**Table 2.** Density estimates (population estimates as numbers of individuals per 100m$^2$) of fishes captured at each site, October 2016.

| | Site 1 | Site 2 | Site 3 | Site 4 | Site 5 | Site 6 |
|---|---|---|---|---|---|---|
| **Species** | **n/100m² ± 95% CL** | **n/100m² ± 95% CL** | **n/100m² ± 95% CL** | **n/100m² ± 95% CL** | **n/100m² ± 95% CL** | **n/100m² ± 95% CL** |
| Bleak | 41.8±24.9 | 3.0±0.6 | 3.0±1.1 | 1.20±0.34 | 0.13±0.00 | 0.00±0.00 |
| Bullhead | 0.0±0.0 | 1.0±0.5 | 0.2±0.1 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| Chub | 1.3±0.1 | 6.7±0.4 | 7.8±2.2 | 0.80±0.27 | 10.00±3.82 | 0.60±0.08 |
| Common Bream | 1.3±0.2 | 0.0±0.0 | 1.0±0.1 | 0.00±0.00 | 0.25±0.00 | 0.00±0.00 |
| Dace | 0.3±0.2 | 0.3±0.3 | 2.2±0.2 | 0.20±0.18 | 0.25±0.16 | 0.00±0.00 |
| Gudgeon | 0.3±0.1 | 1.0±0.4 | 2.0±0.8 | 0.00±0.00 | 1.50±0.59 | 0.00±0.00 |
| Minnow | 2.0±0.7 | 0.0±0.0 | 2.2±0.8 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| Perch | 0.5±0.20 | 4.0±0.2 | 3.2±1.1 | 1.00±0.14 | 0.75±0.08 | 1.40±0.18 |
| Pike | 0.5±0.2 | 0.7±0.4 | 1.0±0.3 | 0.60±0.17 | 2.13±0.96 | 0.80±0.26 |
| Roach | 15.5±8.2 | 17.7±0.5 | 6.8±0.7 | 15.4±0.59 | 5.75±1.41 | 1.20±0.33 |
| Ruffe | 0.0±0.0 | 1.0±0.0 | 0.2±0.1 | 0.00±0.00 | 0.13±0.00 | 0.00±0.00 |
| Tench | 0.0±0.0 | 0.0±0.0 | 0.4±0.2 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| Roach/Bream hybrid | 0.0±0.0 | 0.3±0.2 | 0.0±0.0 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| **Species Richness** | **9** | **9*** | **12** | **6** | **9** | **4** |

**plus roach/bream hybrid*

Length distributions of bleak at site 1 ranged between 22-121 mm, between 94-115 mm at site 2, between 38-123 mm at site 3, between 96-119 mm at site 4 and the one bleak caught at site 5 was 124 mm (Figure 3a). With the exception of site 1, the most common size range of bleak across the

sites was between 80-130 mm. Length distributions of roach at site 1 ranged between 69-248 mm, between 62-183 mm at site 2, between 58-195 mm at site 3, between 55-185 mm at site 4 and between 54-190 mm at site 5 (Figure 3b). The most common size range of roach across all sites was between 90-140 mm. Length distributions of chub at site 1 ranged between 23-445 mm, between 123-179 mm at site 2, between 24-342 mm at site 3, between 123-242 mm at site 4 and between 62-346 mm at site 5 (Figure 4a). The most common size range of chub across all sites was between 140-200 mm; no 0+ chub were captured at site 2 and 4 (Figure 4a). Length distributions of perch at site 1 ranged between 124-159 mm, between 64-208 mm at site 2, between 75-370 mm at site 3, between 73-187 mm at site 4 and between 55-219 mm at site 5. The most common size range of perch across all sites was between 60-220 mm whilst only larger perch (320-380 mm) were captured at Site 3 (Figure 4b). Raw length data are provided in Appendix 1.

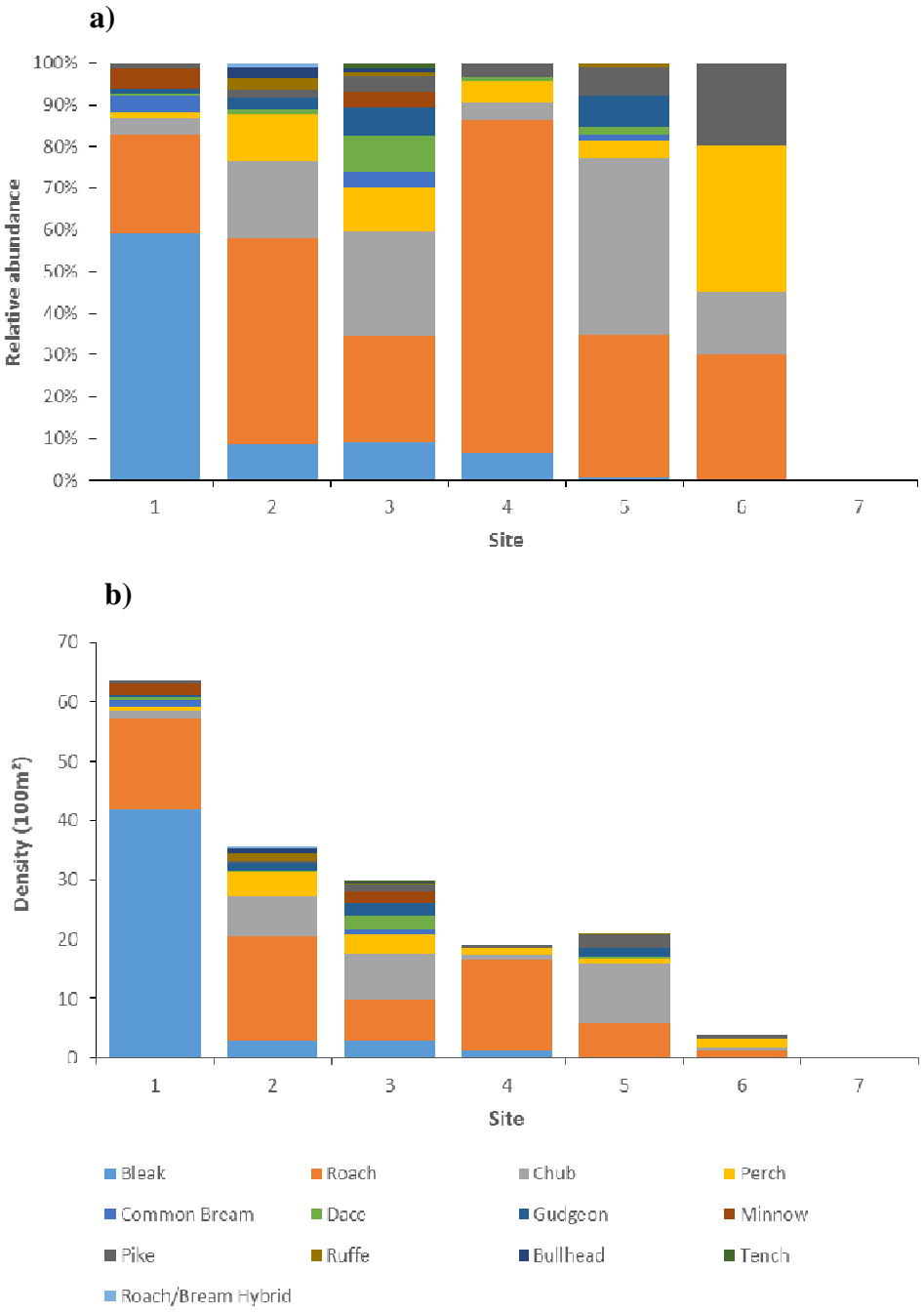


**Figure 2.** Fish species composition (a) and density (b) at six sites west of Oxford, October 2016.

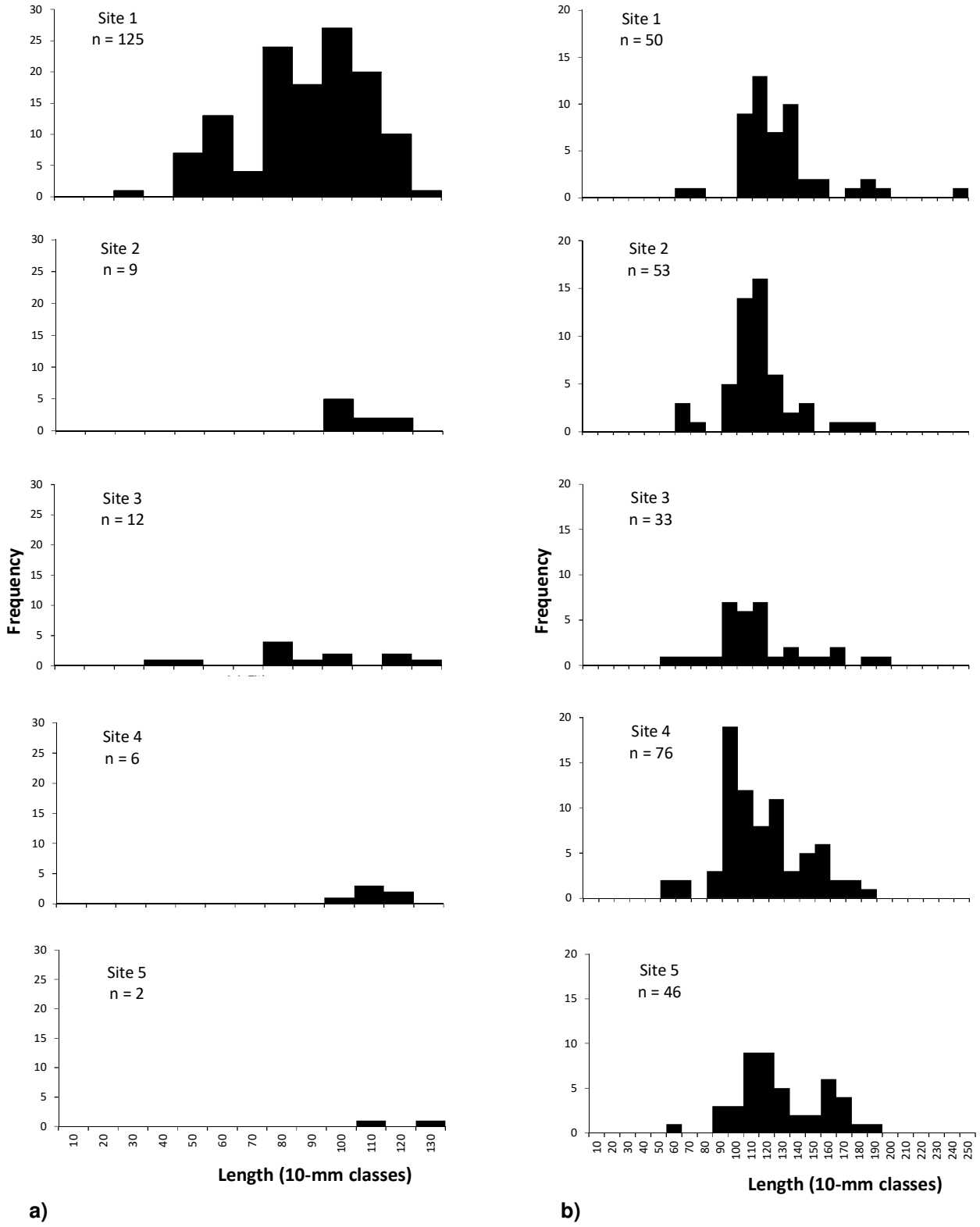


**Figure 3.** Length distribution of bleak (a) and roach (b) at sites 1-5, October 2016.

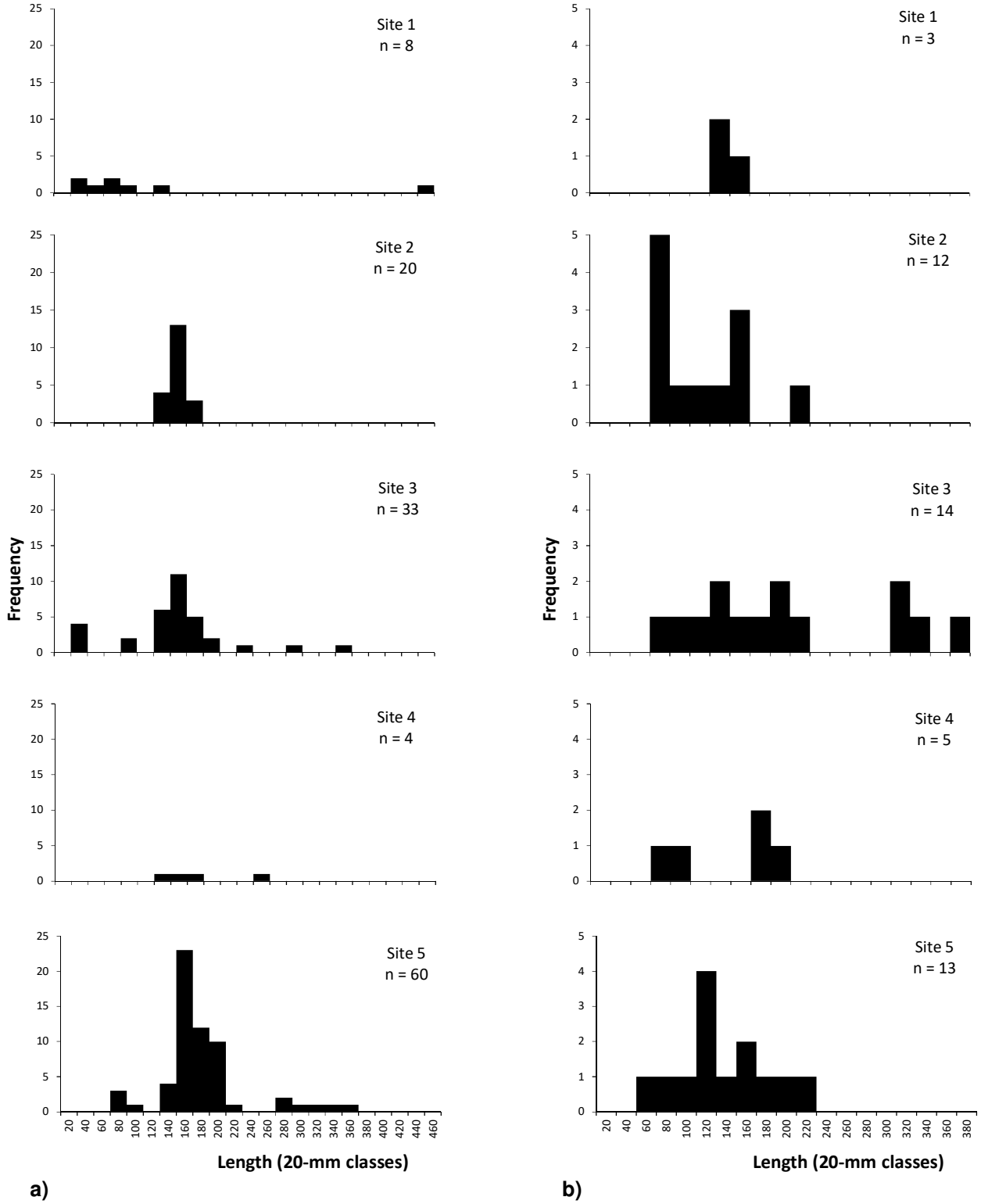


**Figure 4.** Length distribution of chub (a) and perch (b) at sites 1-5, October 2016.

### *3.2 eDNA Results*

#### *3.2.1 Illumina MiSeq sequence runs*

Two libraries were sequenced (12S and CytB) and their run performance is summarised in Table 3.

**Table 3.** Summary of CytB and 12S sequencing runs, including: the specific locus amplified (Locus); the flow cell chemistry (Version), the loading concentration of the library (pM), the percentage PhiX spike-in (% PhiX); the proportion of reads aligned to PhiX (% aligned); the replicate number (Rep); the total reads passing filter (Total reads PF), and the proportion of reads with a quality score of Q30 or above (% Q30).

| Locus | Version | pM | % PhiX | % aligned | Rep | Total reads PF | % Q30 |
|---|---|---|---|---|---|---|---|
| CytB | V3 | 14 | 15 | 22.53 | 1 | 1055406 | 75.4 |
| | | | | | 2 | 1060324 | 76.0 |
| | | | | | 3 | 1035948 | 74.0 |
| 12S | V2 | 13 | 10 | 11.7 | 1 | 811584 | 79.2 |
| | | | | | 2 | 701258 | 78.5 |
| | | | | | 3 | 833666 | 76.9 |

#### 3.*2.2 Taxonomic assignments across Kennington Pond*

Based on the custom curated database and the complete nucleotide database on Genbank, for CytB, 99.1% reads have been taxonomically assigned. The majority of these were assigned to fin-rayed fishes, whereas less than 0.05% reads are assigned to non-fish taxa. For 12S, only 65.38% reads have been taxonomically assigned. The majority of the assigned reads was to fin-rayed fishes (47.60%). The dataset contained a significant amount of non-fish vertebrate sequences from amphibians (0.08%), birds (16.24%) and mammals (1.44%, including humans) (Figure 5).

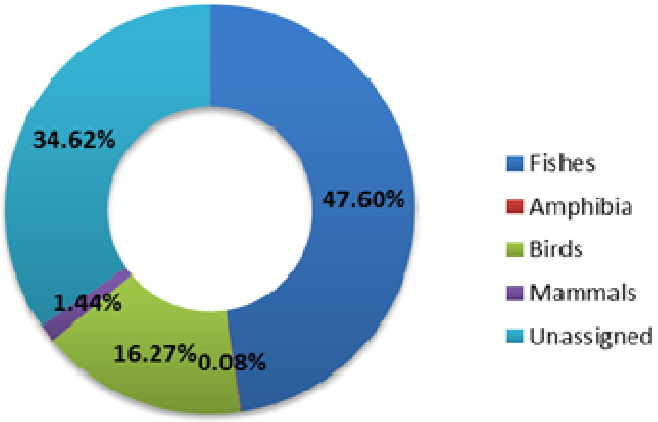


**Figure 5.** The proportional assignments of 12S sequence reads to major taxonomic groups.

*3.2.3 Fish species detection in Kennington Pond*

Across all samples eDNA of ten species was detected: Common Bream, Silver Bream (*Blicca bjoerkna*), Common Carp (*Cyprinus carpio*), Pike, Ide (*Leuciscus idus*), Perch, Roach, Rudd (*Scardinius erythrophthalmus*), Chub and Tench.

The six most common species were detected with both markers (highest confidence of presence: *A. brama*, *B. bjoerkna*, *E. lucius*, *P. fluviatilis*, *R. rutilus*, and *T. tinca*), whereas three species were only detected above the filtering threshold with one of the two markers (CytB: *C. carpio* and *S. erythrophthalmus*; 12S: *S. cephalus*). One species was present at very low read counts with only one marker (12S: *L. idus*) (Figure 6). When applying the filtering criteria, *L. idus* was no longer present and *S. cephalus* was absent for CytB (Table 4 & 5). The highest total number of read counts was recorded for *E. lucius*, *P. fluviatilis* and *R. rutilus* (Figure 7) with a total of over 100,000 reads for at least one of the two markers.

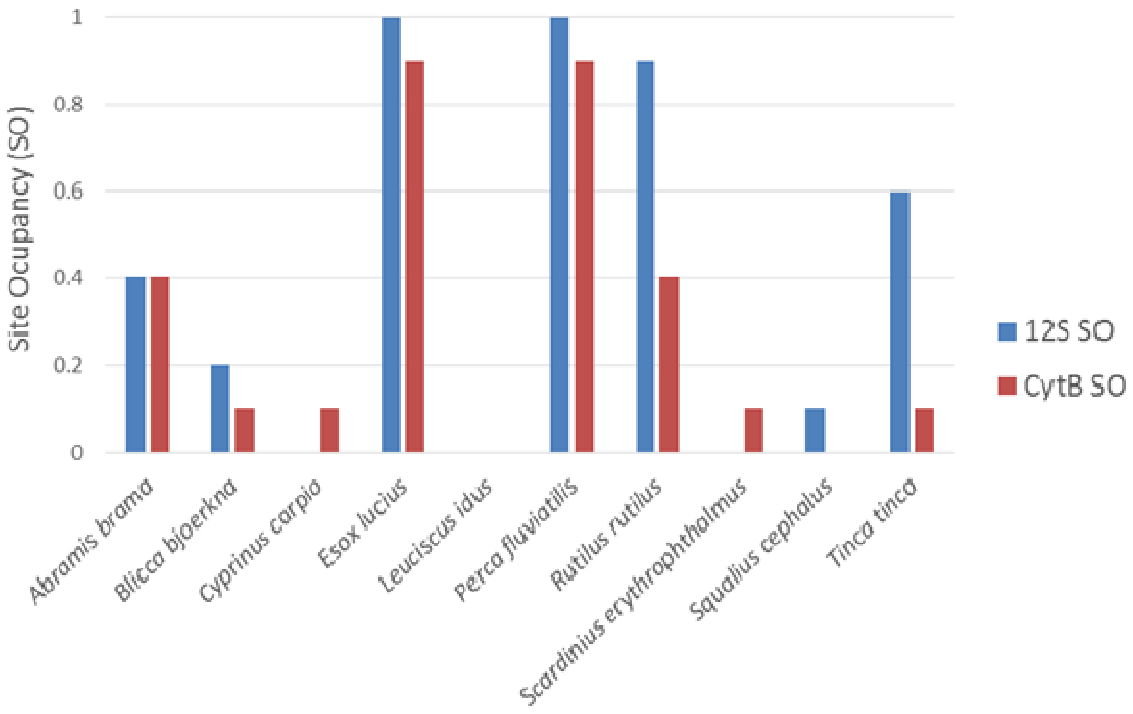


**Figure 6.** The proportion of sample from where the species was found (Site Occupancy, SO) at Kennington Pond for each of the two markers. Star symbols indicate species which were present at very low read counts (below the frequency threshold).

**Table 4.** Total read counts (RC) and site occupancy (SO) for 12S and CytB across all ten samples from Kennington Pond *before* applying filtering criteria based on the custom curated database.

| | 12S RC | CytB RC | 12S SO | CytB SO |
|---|---|---|---|---|
| ***Abramis brama*** | 2474 | 22718 | 0.6 | 0.7 |
| ***Blicca bjoerkna*** | 2706 | 323 | 0.5 | 0.2 |
| ***Cyprinus carpio*** | 0 | 51 | 0 | 0.7 |
| ***Esox lucius*** | 90673 | 308850 | 1 | 1 |
| ***Leuciscus idus*** | 36 | 0 | 0.1 | 0 |
| ***Perca fluviatilis*** | 74658 | 637739 | 1 | 0.9 |
| ***Rutilus rutilus*** | 119376 | 31458 | 1 | 0.5 |
| ***Scardinius erythrophthalmus*** | 0 | 842 | 0 | 0.1 |
| ***Squalius cephalus*** | 3911 | 13 | 0.5 | 0.1 |
| ***Tinca tinca*** | 11047 | 1628 | 1 | 0.3 |
| **Cyprinidae** | 6042 | 343 | 0.8 | 0.2 |
| **Unassigned** | 393056 | 128098 | 1 | 0.8 |

**Table 5.** Total read counts (RC) and site occupancy (SO) for 12S and CytB across all ten samples from Kennington Pond *after* applying filtering criteria based on the custom curated database.

| | 12S RC | CytB RC | 12S SO | CytB SO |
|---|---|---|---|---|
| ***Abramis brama*** | 2466 | 22514 | 0.4 | 0.4 |
| ***Blicca bjoerkna*** | 2688 | 275 | 0.2 | 0.1 |
| ***Cyprinus carpio*** | 0 | 14 | 0 | 0.1 |
| ***Esox lucius*** | 90673 | 308845 | 1 | 0.9 |
| ***Leuciscus idus*** | 0 | 0 | 0 | 0 |
| ***Perca fluviatilis*** | 74658 | 637739 | 1 | 0.9 |
| ***Rutilus rutilus*** | 119367 | 31369 | 0.9 | 0.4 |
| ***Scardinius erythrophthalmus*** | 0 | 842 | 0 | 0.1 |
| ***Squalius cephalus*** | 3877 | 0 | 0.1 | 0 |
| ***Tinca tinca*** | 10896 | 1297 | 0.6 | 0.1 |
| **Cyprinidae** | 6012 | 315 | 0.5 | 0.1 |
| **Unassigned** | 393056 | 128095 | 1 | 0.7 |

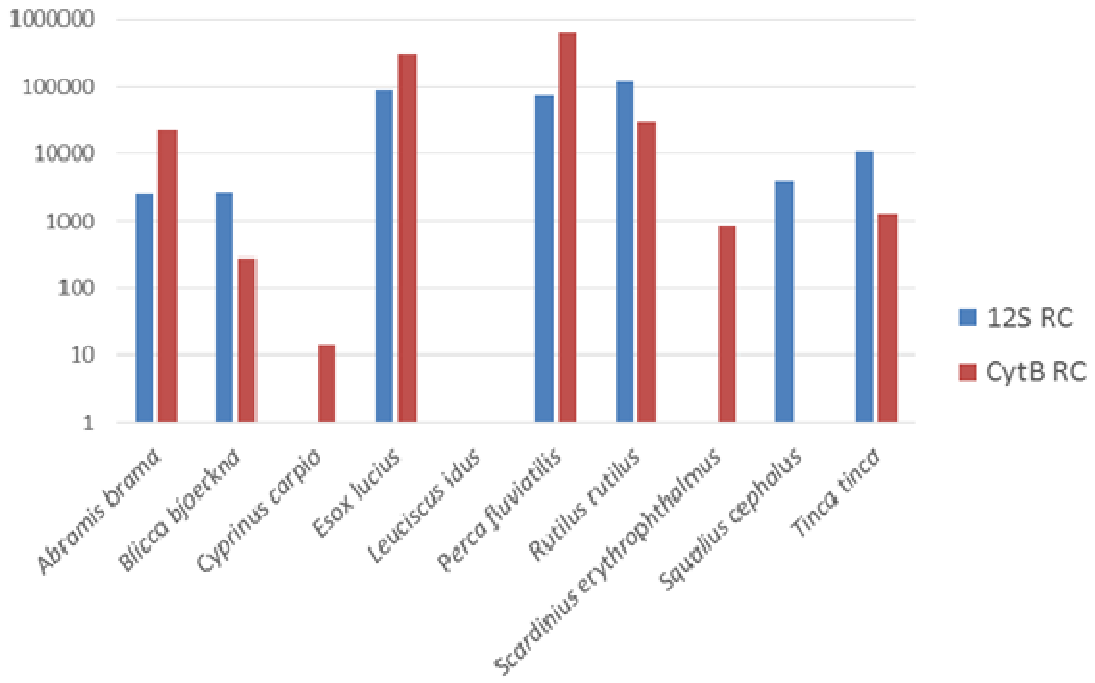


**Figure 7**. Total number of species-specific sequences (read counts, RC) for each of the two markers from three sequencing runs. Note that the y-axis is of logarithmic scale.

## 4. DISCUSSION

### *4.1 Fisheries surveys*

Sites 1 to 6 of Oxford FAS appear to support fish assemblages expected for lowland streams. The assemblages are dominated by eurytopic ('generalist') species, particularly roach, chub, perch and bleak and this was probably due to the availability of suitable habitat such as slow to moderate flows, areas of deep water, bankside vegetation and silt/sand or gravel substrates (Cowx 2001). Roach was consistently in the top 2 most abundant species at sites 1 to 6, which is probably attributable to their adaptability to the available habitats. For example, roach thrive in slow or moderate flow and are phyto-lithophils (spawn on aquatic plants or bed substrate). Which matches available spawning habitat at all sites (Balon 1975; Cowx 2001). Other species that accounted for <10 % of fish captured at a number of sites were common bream, dace, gudgeon, minnow, ruffe, bullhead, tench and roach/bream hybrid. Tench and ruffe are limnophilic (still or slow moving water specialist) species and were present in low numbers at site 2 and 3, where there was low flow and dense areas of instream vegetation. The low number of juvenile fish recorded for all species at all sites could potentially indicate poor recruitment, but it is possible that the densities of juvenile fish were underestimated due to low sampling efficiency by electric fishing because the water was deep and in some areas the river was wide and marginal macrophyte growth was dense.

Site 7 Chilswell drain was extremely overgrown with limited access to the river. The full river length was walked to find a suitable site but only a 20-m stretch was accessible to sample and no fish were captured. Chilswell drain may be more suitable to sample in the spring before the bankside vegetation has grown and it is suggested that the most suitable location is upstream of the road bridge (NGR: SP 51627 03726).

### *4.2 eDNA*

#### *4.2.1 Filtering criteria and level of confidence*

Species identification through metabarcoding needs to account for the presence of false positive records (defined by the detection of a species that is not present in the sample). False positive records can arise from a variety of sources, such as cross-contamination between samples or laboratory contamination, but usually result in records of low frequency. The number of false positives can be reduced by accepting only species records above a certain frequency (filtering threshold, see *Materials and Methods*). Additionally, we have used consistency across three sequencing replicates to reduce the number of false positives caused specifically by barcode misassignments during sequencing. The resulting dataset after applying these criteria should be seen as a conservative result, i.e. the results only include assignments of highest confidence. However, it should be noted that sequences which have not passed this conservative filtering stage may still represent true records; they can simply not be distinguished with high confidence from false positives.

#### *4.2.2 Consistency across markers*

Overall there was a good correlation between markers, both for read counts and site occupancy. Nevertheless there were differences in the relative eDNA abundance for certain species, especially for Roach and Tench, which were recorded at substantially lower abundance for CytB compared to 12S. This pattern is consistent with previous studies and reflects a primer bias in CytB, i.e. these species do not amplify as well during PCR. Therefore the relative abundance estimates based on the 12S marker for these two species are likely to be more representative.

#### *4.2.3 Species detection and abundance*

Our previous study has demonstrated that eDNA metabarcoding can provide qualitative and quantitative information on fish communities in large lakes, outperforming established methods in terms of the number of species detected (Hänfling et al. 2016). Furthermore, recent studies have shown that the method appears to perform well across a wider range of lake environments. Nevertheless, a fully validated eDNA lake classification tool has not yet been developed and the interpretation of eDNA based relative abundance estimates should be guided by the points discussed above, and it might be advisable to use abundance categories rather than exact rankings. It should also be considered that the eDNA based estimates are more likely to reflect biomass rather than species counts. Taking such an approach would classify Perch, Roach and Pike as highly abundant in Kennington Pond; Common Bream and Tench as abundant; Silver Bream as rare; Common Carp, Rudd and Chub as very rare (with a possibility that these three records represent a false positive); and Ide as a possible occurrence at extremely low density.

## 5. REFERENCES


Balon E.K. (1975). Reproductive guilds of fishes: a proposal and definition. *Journal of the Fisheries Research Board of Canada*. **32**, 821-864.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30 (15), 2114-2120.

Carle, F.L. and Strub, M.R. (1978). A new method for estimating population size from removal data. *Biometrics* **34**, 621-830.

Cowx, I. G. (2001). *Factors Influencing Coarse Fish Populations in Rivers*. R&D Publication 18. Environment Agency, Bristol.

Hänfling, B., Lawson Handley, L., Read, D.S., Hahn, C., LI, J., Nichold, P., Blackman, R.C., Oliver, A. & Winfield, I.J. 2016. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular Ecology, 25, 3101-3119.

Huson, D.H., Auch, A.F., QI, J. & Schuster, S.C. 2007. MEGAN analysis of metagenomic data. Genome Research, 17, 377-386.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X. & Wilson, A.C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences of the United States of America, 86, 6196-6200.

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

Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. 2000. A Greedy Algorithm for Aligning DNA Sequences. Journal of Computational Biology, 7, 203-214.

Zippin C. (1956). An evaluation of the removal method for estimating animal populations. *Biometrics,* **8**,163-189.

**Plate 1.** Site 1 Seacourt Stream, October 2016.

**Plate 2.** Site 2 Seacourt Stream by Park and Ride, October 2016.

**Plate 3.** Site 3 Botley Stream by Golf Range, October 2016.

**Plate 4.** Site 4 Seacourt Stream at Hinksey, October 2016.

**Plate 5.** Site 5 Bulstake Stream, October 2016.

**Plate 6.** Site 6 Hinksey Stream, October 2016.

**Plate 7**. Site 7 Chilswell drain, October 2016.

**Plate 8**. Kennington Pond, October 2016.

# Appendix 1.

Lengths (mm) of all fishes captured per run at sites 1-6 (no fish were caught at site 7), October 2016.

**Site 1 – Seacourt stream**

| | Common bream | Pike | Bleak | | | | | Roach | | | Perch | Chub | Minnow | Dace | Gudgeon |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Run 1 | 362 | 554 | 88 | 78 | 73 | 90 | 53 | 109 | 111 | | 127 | 138 | 30 | 48 | 121 |
| | 366 | | 75 | 100 | 115 | 113 | 74 | 132 | 71 | | | 57 | 37 | | |
| | 159 | | 76 | 106 | 76 | 101 | 75 | 120 | 116 | | | 23 | 38 | | |
| | 118 | | 64 | 54 | 102 | 101 | 81 | 135 | | | | 70 | | | |
| | 117 | | 121 | 103 | 103 | 91 | 95 | 138 | | | | 83 | | | |
| | | | 114 | 97 | 75 | 57 | 90 | 123 | | | | | | | |
| | | | 66 | 98 | 103 | 75 | 53 | 106 | | | | | | | |
| | | | 91 | 92 | 80 | 87 | 49 | 69 | | | | | | | |
| | | | 103 | 104 | 91 | 98 | 99 | 114 | | | | | | | |
| | | | 116 | 92 | 62 | 50 | 78 | 108 | | | | | | | |
| Run 2 | 162 | 640 | 118 | 112 | 89 | 45 | | 108 | 122 | 149 | 159 | 445 | 55 | | 72 |
| | 118 | 620 | 110 | 112 | 54 | 102 | | 106 | 110 | 116 | | 63 | 25 | | |
| | | | 48 | 90 | 94 | 58 | | 128 | 134 | 115 | | 28 | 29 | | |
| | | | 88 | 72 | 22 | 88 | | 119 | 198 | 101 | | | 48 | | |
| | | | 97 | 51 | 85 | 94 | | 140 | 248 | 108 | | | 50 | | |
| | | | 79 | 67 | 119 | 93 | | 111 | 130 | 125 | | | | | |
| | | | 107 | 103 | 76 | 85 | | 115 | 133 | 115 | | | | | |
| | | | 103 | 74 | 47 | 100 | | 177 | 124 | 153 | | | | | |
| | | | 104 | 48 | 55 | 100 | | 185 | 134 | | | | | | |
| | | | 76 | 51 | 47 | | | 119 | 120 | | | | | | |
| Run 3 | 319 | | 96 | 74 | 103 | 75 | | 144 | | | 124 | | 34 | 47 | |
| | | | 87 | 85 | 102 | 110 | | 138 | | | | | 30 | | |
| | | | 77 | 88 | 97 | 75 | | 135 | | | | | | | |
| | | | 82 | 98 | 58 | 97 | | 158 | | | | | | | |
| | | | 98 | 98 | 74 | 58 | | 102 | | | | | | | |
| | | | 72 | 102 | 105 | 78 | | 185 | | | | | | | |
| | | | 57 | 87 | 98 | | | 139 | | | | | | | |
| | | | 58 | 99 | 112 | | | 116 | | | | | | | |
| | | | 92 | 82 | 100 | | | 128 | | | | | | | |
| | | | 119 | 76 | 82 | | | | | | | | | | |

**Site 2 – Seacourt stream park & ride**

| | Pike | Bleak | Roach | | | | | Perch | Chub | | Dace | Gudgeon | Ruffe | Bullhead | Ro/Br hybrid |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Run 1 | | 113 | 180 | 117 | 128 | 116 | 120 | 208 | 151 | 147 | | | 138 | | |
| | | 100 | 170 | 102 | 95 | 110 | | 145 | 128 | 148 | | | 134 | | |
| | | 97 | 111 | 130 | 146 | 109 | | 156 | 145 | 158 | | | 128 | | |
| | | 105 | 132 | 117 | 142 | 106 | | 147 | 146 | 160 | | | | | |
| | | | 110 | 106 | 117 | 110 | | 72 | 176 | 123 | | | | | |
| | | | 110 | 108 | 120 | 123 | | 138 | 160 | | | | | | |
| | | | 132 | 110 | 112 | 111 | | 100 | 151 | | | | | | |
| | | | 117 | 69 | 118 | 68 | | 64 | 134 | | | | | | |
| | | | 114 | 115 | 95 | 113 | | 72 | 170 | | | | | | |
| | | | 125 | 107 | 98 | 108 | | 68 | 147 | | | | | | |
| Run 2 | 152 | 100 | 122 | | | | | 68 | 179 | | | 81 | | 50 | 140 |
| | | 103 | 183 | | | | | | 155 | | | 61 | | 50 | |
| | | 100 | 109 | | | | | | 143 | | | 96 | | | |
| | | 94 | 109 | | | | | | 140 | | | | | | |
| | | | 93 | | | | | | | | | | | | |
| | | | 117 | | | | | | | | | | | | |
| | | | 118 | | | | | | | | | | | | |
| | | | 98 | | | | | | | | | | | | |
| | | | 62 | | | | | | | | | | | | |
| | | | 71 | | | | | | | | | | | | |
| Run 3 | 120 | 115 | 142 | | | | | 110 | 150 | | 106 | | | 47 | |
| | | | 126 | | | | | | | | | | | | |

**Site 3 – Botley stream by Golf Range**

| | Common bream | Pike | Bleak | Roach | | | Perch | Chub | | Dace | Gudgeon | Tench | Ruffe | Bullhead | Minnow |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Run 1 | 130 | 170 | 75 | 104 | 116 | 96 | 196 | 286 | 222 | 146 | 98 | | | | |
| | 86 | | 90 | 105 | 113 | 139 | 340 | 342 | 141 | 136 | 101 | | | | |
| | 103 | | 111 | 97 | 112 | | 130 | 144 | 143 | 144 | 90 | | | | |
| | | | | 99 | 134 | | 139 | 163 | 135 | 160 | | | | | |
| | | | | 104 | 94 | | 75 | 186 | 163 | 125 | | | | | |
| | | | | 110 | 107 | | | 172 | 135 | 136 | | | | | |
| | | | | 114 | 161 | | | 135 | 144 | 138 | | | | | |
| | | | | 58 | 166 | | | 140 | | 128 | | | | | |
| | | | | 151 | 128 | | | 138 | | | | | | | |
| | | | | 114 | 97 | | | 145 | | | | | | | |
| Run 2 | 106 | 176 | 123 | 142 | | | 310 | 155 | | 116 | 173 | 200 | 130 | 60 | |
| | 97 | 348 | 74 | 195 | | | 370 | 97 | | 140 | 110 | | | | |
| | | 172 | 75 | 82 | | | 102 | 149 | | | 92 | | | | |
| | | | 111 | 74 | | | 150 | 170 | | | 53 | | | | |
| | | | 98 | 105 | | | 213 | 145 | | | | | | | |
| | | | 44 | 92 | | | 175 | 200 | | | | | | | |
| | | | 100 | 182 | | | 98 | 157 | | | | | | | |
| | | | | | | | | 144 | | | | | | | |
| | | | | | | | | 174 | | | | | | | |
| Run 3 | | 325 | 72 | 98 | | | 188 | 122 | | 144 | 68 | 135 | | | 35 |
| | | | 38 | 112 | | | 310 | 158 | | | 100 | | | | 45 |
| | | | | 112 | | | | 100 | | | | | | | 42 |
| | | | | 65 | | | | 32 | | | | | | | 31 |
| | | | | | | | | 24 | | | | | | | 27 |
| | | | | | | | | 28 | | | | | | | |
| | | | | | | | | 25 | | | | | | | |

**Site 4 – Seacourt stream at North Hinksey**

| | Pike | Bleak | Roach | | | | | | Perch | Chub | Dace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Run 1 | 502 | 110 | 99 | 64 | 97 | 96 | 101 | 126 | 174 | 161 | |
| | 160 | 119 | 99 | 62 | 116 | 98 | 101 | 113 | 84 | 123 | |
| | | | 98 | 90 | 144 | 119 | 123 | 148 | 73 | | |
| | | | 103 | 97 | 104 | 99 | 130 | 97 | | | |
| | | | 103 | 98 | 151 | 55 | 137 | 121 | | | |
| | | | 97 | 99 | 113 | 87 | 163 | 141 | | | |
| | | | 107 | 57 | 111 | 185 | 156 | | | | |
| | | | 113 | 94 | 128 | 151 | 109 | | | | |
| | | | 100 | 141 | 101 | 158 | 152 | | | | |
| | | | 98 | 103 | 100 | 128 | 140 | | | | |
| Run 2 | | 103 | 98 | 152 | | | | | 163 | | |
| | | 110 | 128 | 86 | | | | | 187 | | |
| | | 96 | 100 | 172 | | | | | | | |
| | | | 106 | 172 | | | | | | | |
| | | | 142 | 130 | | | | | | | |
| | | | 123 | 103 | | | | | | | |
| | | | 129 | | | | | | | | |
| | | | 111 | | | | | | | | |
| | | | 104 | | | | | | | | |
| | | | 119 | | | | | | | | |
| Run 3 | 179 | 118 | 127 | | | | | | | 149 | 109 |
| | | | 164 | | | | | | | 242 | |
| | | | 100 | | | | | | | | |
| | | | 132 | | | | | | | | |

**Site 5 – Bulstake Stream**

| | Common bream | Pike | Bleak | Roach | | Perch | Chub | | | | Dace | Gudgeon | Ruffe |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Run 1 | 119 | 300 | 124 | 103 | 97 | 113 | 72 | 162 | | | | 92 | 121 |
| | 161 | 186 | | 115 | 167 | 181 | 142 | 346 | | | | 104 | |
| | | | | 167 | 102 | 120 | 62 | 317 | | | | 83 | |
| | | | | 100 | 159 | 108 | 165 | 158 | | | | 80 | |
| | | | | 165 | 144 | 174 | 200 | 261 | | | | | |
| | | | | 134 | 132 | | 192 | 164 | | | | | |
| | | | | 112 | 120 | | 300 | 210 | | | | | |
| | | | | 108 | 143 | | 195 | | | | | | |
| | | | | 158 | | | 182 | | | | | | |
| | | | | 190 | | | 94 | | | | | | |
| Run 2 | | 283 | | 123 | 93 | | 159 | 170 | 138 | 152 | | 81 | |
| | | 214 | | 159 | 178 | | 151 | 160 | 272 | 175 | | 78 | |
| | | 156 | | 109 | 129 | | 185 | 184 | 168 | 121 | | 93 | |
| | | 194 | | 121 | 126 | | 182 | 159 | 146 | 142 | | | |
| | | | | 102 | 118 | | 165 | 160 | 155 | 143 | | | |
| | | | | 158 | 156 | | 149 | 186 | 152 | | | | |
| | | | | 159 | 89 | | 180 | 150 | 160 | | | | |
| | | | | 162 | 87 | | 180 | 139 | 185 | | | | |
| | | | | 118 | | | 163 | 148 | 155 | | | | |
| | | | | 106 | | | 160 | 332 | 156 | | | | |
| Run 3 | | 186 | | 54 | | 132 | 69 | | | | 149 | 101 | |
| | | 215 | | 104 | | | 142 | | | | 149 | 85 | |
| | | 169 | | 107 | | | 121 | | | | | 82 | |
| | | 145 | | 118 | | | 158 | | | | | | |
| | | | | | | | 166 | | | | | | |

**Site 6 – Hinksey stream**

| | Pike | Roach | Perch | Chub |
|---|---|---|---|---|
| Run 1 | 151 | 104 | 219 | 191 |
| | | 119 | 160 | 146 |
| | | | 144 | |
| | | | 81 | |
| Run 2 | 133 | 111 | 80 | 170 |
| | 122 | 130 | 120 | |
| | | 118 | 55 | |
| Run 3 | 147 | 88 | | |