Environmental DNA (eDNA) Biomonitoring Project Report

Item	Details
Report title	eDNA Metabarcoding Survey
Project ID	25010_01
Client	Challenging Habitat for Pelican of London Ltd
Client contact	Charlotte Braungardt
Project location	Plymouth Sound National Marine Park / English Channel
Sector	Education / Conservation
Date of report	05 January 2026
Prepared by	Applied Genomics Ltd
Reviewed by	Applied Genomics Ltd

Applied Genomics Ltd 09180742 | VAT: GB903660148
Registered Office: Asteroceras Science Park | Unit C | Greengates Way | Norfolk | NR12 8ED
: 01803 611818 | : https://appliedgenomics.co.uk | @: solutions@appliedgenomics.co.uk

EXECUTIVE SUMMARY

1.1 Study purpose and context

Challenging Habitat commissioned this environmental DNA (eDNA) metabarcoding study on behalf of Pelican of London Ltd to support educational sail-training programmes aimed at A-level students. The objective was to demonstrate how water sampling and molecular methods can be used to detect vertebrate biodiversity in marine environments, while generating a baseline dataset that could be expanded into a time series through repeat sampling.

1.2 What was done

Two high-volume water samples were collected during a September 2025 voyage in southern England, including Plymouth Sound National Marine Park and a second location along the English Channel. Samples were collected using inDepth water samplers over a full tidal cycle and analysed using a vertebrate-targeted 16S rRNA metabarcoding marker. Laboratory processing, sequencing and bioinformatic analysis followed standard Applied Genomics workflows, with appropriate negative controls included throughout.

1.3 Key findings

Vertebrate DNA was successfully detected in both samples, with taxonomic assignments dominated by marine fish and other chordates expected in coastal waters of southern England. Multiple taxa were detected at genus and species level, with several taxa represented by more than one haplotype, indicating detectable within-taxon genetic diversity. Negative controls showed low read counts overall, supporting confidence that the reported detections primarily reflect true environmental signals.

1.4 Confidence, limitations and implications

The results provide a robust qualitative snapshot of vertebrate biodiversity present in the sampled water masses at the time of sampling. As with all metabarcoding surveys, results should be interpreted primarily as presence/absence and relative patterns rather than precise measures of abundance. The limited number of samples constrains spatial and temporal inference, but the dataset provides a strong foundation for future repeat sampling, comparative analyses and educational interpretation.

2. PROJECT OVERVIEW AND SCOPE

2.1 Background

Pelican of London Ltd is a not-for-profit sail training charity delivering educational programmes focused on marine science and environmental monitoring. This project forms

part of a broader aim to introduce students to modern biodiversity assessment techniques and to build a repeatable monitoring framework over time.

2.2 Objectives

1. Demonstrate the application of eDNA metabarcoding for detecting marine vertebrate biodiversity.
2. Characterise vertebrate taxa present at two coastal marine locations during the September 2025 voyage.
3. Generate baseline data suitable for future time-series comparison and educational use.

2.3 Study design overview

Two inDepth water samples were collected, each representing an integrated sample over approximately one full tidal cycle (~12 hours). Standard field, extraction and PCR negative controls were included to monitor contamination.

Table 1. Summary of study design and sample types.

Sample type	Description	Number	Notes
Water (inDepth)	High-volume integrated water samples	2	Full tidal cycle
Extraction negatives	Blank extractions	2	Contamination control
PCR negatives	No-template controls	2	Contamination control

2.4 Genetic markers and target groups

Table 2. Markers and target biological groups.

Target group	Gene marker	Marker name	Application
Vertebrates	16S rRNA	16S MV3	Fish and other vertebrates

3. METHODS

3.1 Field sampling

Water samples were collected by the client using an Applied Genomics inDepth water sampler, deployed for approximately 12 hours to integrate eDNA across a full tidal cycle. Filters and preservatives were retained and returned to the Applied Genomics laboratory under controlled conditions.

3.2 Laboratory processing and sequencing

DNA was extracted using protocols optimised for environmental water samples. DNA concentration and quality were assessed prior to amplification. Target 16S rRNA regions were amplified in triplicate PCR reactions per sample. Amplicons were pooled, indexed and sequenced on an Illumina MiSeq platform using paired-end chemistry. Extraction and PCR negative controls were processed alongside samples.

3.3 Bioinformatic processing and taxonomic assignment

Raw sequence reads were quality filtered, denoised and inferred into amplicon sequence variants (ASVs) using the DADA2 algorithm. Chimeric sequences were removed, and negative controls were used to identify and exclude likely contaminants. Taxonomic assignment was performed against curated vertebrate reference databases, with confidence scores retained at each taxonomic rank. Read counts were interpreted as semi-quantitative indicators only.

3.4 Haplotype characterisation and detection credibility

ASV presence/absence data were aggregated by taxon to infer haplotypes. The number of haplotypes per taxon was used as an index of sampled genetic diversity. A Bayesian credibility framework integrated assignment confidence and contextual information to classify detections as Low, Moderate or High credibility.

4. RESULTS

4.1 Data quality and sequencing performance

Both samples yielded high quality amplifiable DNA, and sequencing generated sufficient reads to support vertebrate detection. Negative controls showed low read counts relative to environmental samples.

Applied Genomics Ltd 09180742 | VAT: GB903660148
Registered Office: Asteroceras Science Park | Unit C | Greengates Way | Norfolk | NR12 8ED
: 01803 611818 | : https://appliedgenomics.co.uk | @: solutions@appliedgenomics.co.uk

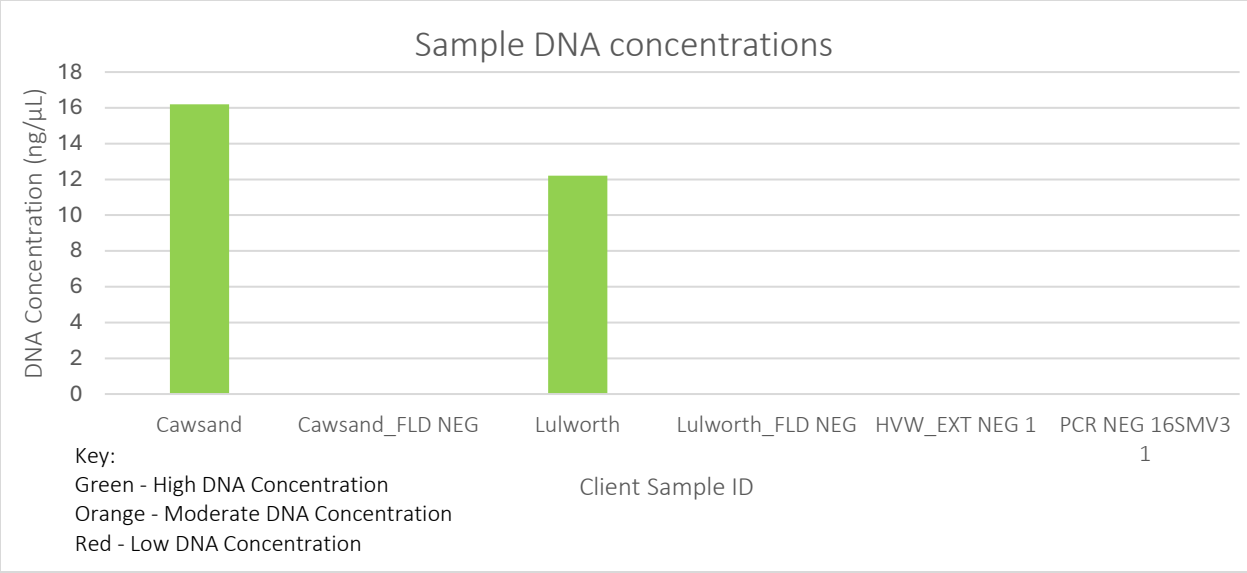

Figure 1: Purified sample DNA concentrations. Samples with sufficient DNA concentration are shown in green.

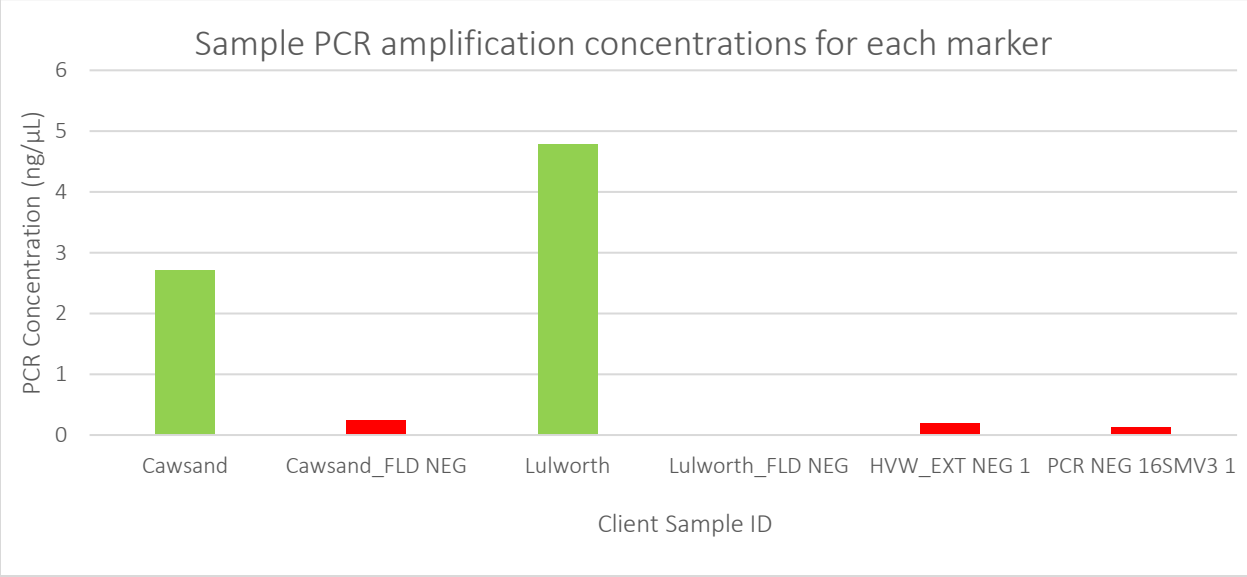

Figure 2: Amplified DNA concentrations for the 16S vertebrate marker.

Applied Genomics Ltd 09180742 | VAT: GB903660148
Registered Office: Asteroceras Science Park | Unit C | Greengates Way | Norfolk | NR12 8ED
: 01803 611818 | : https://appliedgenomics.co.uk | @: solutions@appliedgenomics.co.uk

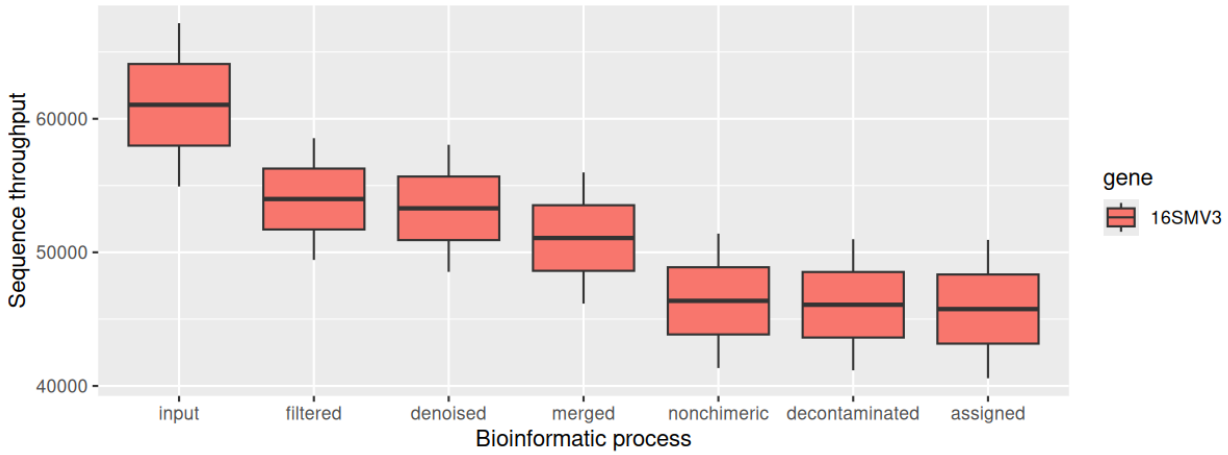

Figure 3: Range of DNA sequences retained in bioinformatics processing, from input sequences through to taxonomic assignment.

4.2 Taxonomic coverage and diversity

Across both samples, multiple vertebrate taxa were detected, predominantly within marine chordates. Genus-level assignments were generally more robust than species-level assignments, consistent with expected marker performance.

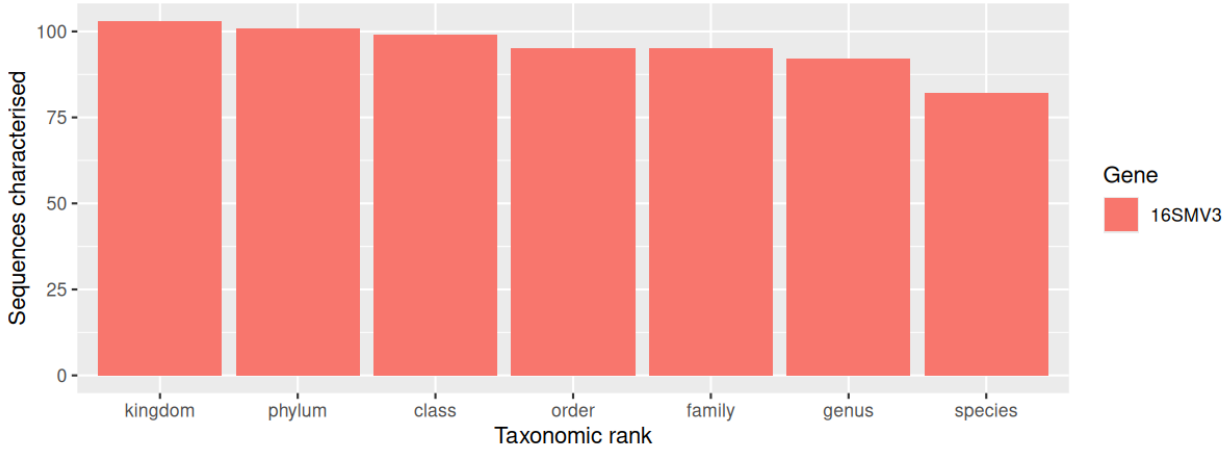

Figure 4: Number of sequences successfully characterised to each taxonomic rank.

4.3 Species and genera of particular interest

The vertebrate 16S metabarcoding dataset included several taxa that are widely regarded as charismatic, ecologically important, or conservation relevant within UK coastal and shelf waters. These taxa are listed below with their common names and reported detection credibility scores, based on the haplotype analysis.

Detected taxa of particular interest include:

- **Conger conger** (European conger eel) – *High credibility*. A large-bodied demersal predator associated with rocky reefs and complex seabed habitats. Multiple haplotypes were detected, supporting a robust signal of presence.
- **Sardina pilchardus** (European sardine / pilchard) – *High credibility*. A schooling pelagic species of high ecological and socio-economic importance, forming a key trophic link between plankton and higher predators.
- **Merluccius merluccius** (European hake) – *Moderate credibility*. An ecologically important demersal predator widely distributed in the northeast Atlantic and English Channel.
- **Atherina presbyter** (Sand smelt) – *Moderate credibility*. A small coastal fish commonly associated with inshore and estuarine environments, and frequently used as an indicator of nearshore habitat use.
- **Labrus** spp. (Wrasse) – *Moderate credibility*. A genus of reef-associated fish that are familiar and recognisable to the public and important components of rocky shore and reef ecosystems.

These detections are consistent with expectations for productive coastal waters such as Plymouth Sound and the wider English Channel. Where High credibility scores were assigned, detections can be interpreted as strong evidence that these taxa occur in, or regularly utilise, the sampled areas. Moderate credibility detections should be interpreted with appropriate caution but nonetheless provide informative signals of likely presence.

Overall, the presence of these charismatic and ecologically significant taxa demonstrates the sensitivity of the applied eDNA approach and its value for both educational engagement and baseline biodiversity assessment.

4.4 Genetic diversity indicators (haplotypes)

Multiple haplotypes were observed for several taxa, suggesting detectable within-taxon genetic diversity. These patterns should be interpreted cautiously and as indicative rather than quantitative.

---

5. DISCUSSION

5.1 Interpretation against project objectives

The study successfully demonstrated the use of eDNA metabarcoding to detect marine vertebrate biodiversity, meeting the project's educational and baseline objectives.

5.2 Ecological and conservation context of key detections

The detection of large-bodied, mobile and ecologically influential fish species highlights the ability of eDNA metabarcoding to capture signals from taxa that may be difficult to observe directly during short field visits. Species such as European conger eel, European sardine and European hake play important roles in coastal and shelf food webs, and their detection provides contextual evidence that the sampled water masses are connected to functioning marine ecosystems spanning both pelagic and demersal habitats.

From an educational perspective, the detection of recognisable and well-known species adds substantial value, allowing students to directly link molecular data to familiar marine organisms and to broader discussions of marine ecology, fisheries and conservation. From a monitoring perspective, these results illustrate how eDNA can complement traditional survey methods by providing broad taxonomic coverage with minimal field effort.

5.3 Methodological considerations and limitations

Key limitations include the small number of samples and the inherent presence/absence nature of metabarcoding data. eDNA transport and persistence may also influence detections.

REFERENCES

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10–12.

McMurdie, P. J., and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, 8(4), e61217.

R Core Team. (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For biodiversity research and monitoring. In *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Valsecchi, E., Bylemans, J., Goodman, S. J., Lombardi, R., Carr, I., Castellano, L., Galimberti, A., & Galli, P. (2020). Novel universal primers for metabarcoding environmental DNA surveys of marine mammals and other marine vertebrates. *Environmental DNA*, *2*(4), 460–476. https://doi.org/10.1002/edn3.72

Wright, E. S. (2016). Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *The R Journal*, 8, 352.

CONFIDENTIALITY AND COPYRIGHT

This report contains confidential information prepared solely for Carcinus Ltd. by Applied Genomics Ltd. The information provided in this document should be treated as commercial-in-confidence. The recipient agrees to hold all information within as confidential and not to use or disclose it to unauthorised parties, in whole or in part, without prior written consent from Applied Genomics Ltd. This obligation applies before, during and after any agreement relating to the services described in this report..

TERMS AND CONDITIONS

Applied Genomics standard Terms and Conditions are published on our website and may be accessed through the following link: https://appliedgenomics.co.uk/terms.

APPENDICES

Appendix A. Data outputs provided

The accompanying data file [placeholder for Data Output workbook] incorporates the following outputs:

1. **Summary** – A summary of the project details and the results presented.

2. **Metadata** – Sample metadata includes unique sample identifiers, collection locations with geographic coordinates, laboratory processing results, genetic marker amplification outcomes, and, where available, associated environmental parameters.

3. **eDNA Full Sequence Dataset** – Contains all assigned amplicon sequence variants (ASVs) with taxonomic ranks, taxonomic assignment confidence values and targeted marker genes. Sequence reads are shown for each sample column for each ASV, along with extraction (EXT NEG) and amplification negative (PCR NEG) controls.

4. **eDNA Full Taxonomy Dataset** - Provides a summary of the eDNA Full Sequence Dataset merging each unique taxonomic rank that could be assigned to the lowest taxonomic level. The numbers simply indicate presence or absence of the taxa at the assigned rank in each sample. Negative controls are excluded.

5. **Credibility Analysis** – This analysis employs risk-based naïve Bayesian methods to assess the likelihood of false detection or misclassification for each haplotype, using historical data from the Global Biodiversity Information Facility (GBIF). The *'GBIF'* column indicates whether a taxon has been previously detected in the area. The *'risk'* column shows the probability of a false-positive detection, while the *'score'* column reflects the detection credibility of each taxon. A higher risk corresponds to lower credibility, while a lower risk indicates higher credibility for the detected taxon.

6. **Haplotypes with population status** – Shows the number of haplotypes (ASVs identified to genus or species) detected for each taxon, along with their taxonomic ranks, and false-detection credibility score.

Applied Genomics Ltd 09180742 | VAT: GB903660148
Registered Office: Asteroceras Science Park | Unit C | Greengates Way | Norfolk | NR12 8ED
: 01803 611818 | : https://appliedgenomics.co.uk | @: solutions@appliedgenomics.co.uk