eDNA Processing and Analysis: From Sequences to Insight

July 2025

Contents

Introduction

After eDNA is collected, it's processed in the lab and analyzed through a series of molecular and computational steps. This section introduces the full pipeline from extraction to ecological insight.

DNA Extraction

Methods

- Commercial kits (e.g., Qiagen, Zymo)
- CTAB protocol (especially for difficult substrates)
- Mechanical lysis (bead beating) for tough samples

Museum Considerations

- Ethanol extracts may require extra clean-up
- Small sample volumes and degraded DNA need gentle handling

PCR Amplification

Key Components

- Target gene (COI, 16S, 18S, ITS)
- Forward and reverse primers with sample-specific barcodes
- Polymerase mix, buffers, thermal cycler

Multiplexing

- Add indexes (barcodes) to each sample
- Allows pooling of hundreds of samples for sequencing

Library Preparation & Sequencing

Platforms

- **Illumina MiSeq**: Short reads, high accuracy, widely used
- **Oxford Nanopore**: Long reads, portable, real-time
- **PacBio**: Long, accurate reads (more expensive)

Library Steps

- Clean and quantify PCR products
- Pool amplicons
- Add adapters for platform-specific sequencing

Bioinformatics Workflow

Quality Control

- Trim adapters, remove low-quality reads
- Filter out chimeras (artifacts)

Sequence Inference

- Denoising tools (e.g., DADA2, UNOISE)
- Generate ASV table (amplicon sequence variants)

Taxonomic Assignment

- Compare ASVs to reference databases (SILVA, MIDORI, BOLD)
- Use BLAST, RDP classifier, or QIIME's built-in classifiers

Diversity and Community Analysis

Alpha Diversity

- Measures richness within a sample
- Shannon, Simpson, Observed ASVs

Beta Diversity

- Compares composition between samples
- Dissimilarity indices (Bray-Curtis, Jaccard)
- Ordination techniques (NMDS, PCA, PCoA)

Visualization and Interpretation

Common Tools

- `phyloseq` in R
- Bar plots, heatmaps, diversity plots
- Network analysis for co-occurrence

Museum Applications

- Archive eDNA extracts with specimen metadata
- Link modern and historical biodiversity
- Contribute to public reference libraries (e.g., GenBank, BOLD)

Challenges and Caveats

- **False positives/negatives** from contamination or low biomass
- **Transport effects**: DNA from upstream or other locations
- **PCR bias**: Not all taxa amplify equally

Preservation and DNA Recovery: Formalin and Ethanol

Why this matters:

- Many museum collections contain valuable specimens preserved decades ago.

- Modern methods like eDNA offer a chance to **recover biodiversity signals** from these specimens.
- Success depends on how the material was stored and preserved.

DNA from Ethanol

- DNA can leach into ethanol from tissue over time.
- Ethanol can be extracted directly and filtered like a water sample.
- Be aware of:
  - Evaporation
  - Microbial contamination
  - Loss of long DNA fragments

DNA from Formalin

- Formalin causes **extensive crosslinking**, degrading DNA quality.
- Extraction success depends on:
  - Sample age
  - Buffering (neutral-buffered formalin is better)
  - Use of reversal techniques (e.g., heat, alkaline lysis, long incubation)

Tips for Success

- Use high-sensitivity kits or protocols for low-input DNA
- Consider using **PCR-free** library prep if DNA is heavily damaged
- Focus on short amplicons (100–200 bp) for degraded samples

Foundational Reading

- **Campos & Gilbert (2012)** – DNA extraction from formalin-fixed samples
  Methods in Molecular Biology, 840, 81–89.
  https://doi.org/10.1007/978-1-61779-516-9_11

- **Hykin et al. (2015)** – Resurrection of historical specimens using genomic approaches
  PeerJ, 3:e967.
  https://doi.org/10.7717/peerj.967

- **Ruane & Austin (2017)** – Unlocking preserved DNA from formalin-fixed herpetological specimens
  PLOS ONE, 12(3): e0173141.
  https://doi.org/10.1371/journal.pone.0173141

- **Nagano et al. (2013)** – DNA damage in ethanol-preserved samples
  Journal of Forensic Sciences, 58(5), 1173–1179.
  https://doi.org/10.1111/1556-4029.12192

Final Thoughts

eDNA analysis is a powerful tool, but it requires rigorous protocols and thoughtful interpretation. Used well, it offers scalable, reproducible insights across fields—from conservation to curation.