# Environmental DNA(eDNA) as a next-generation biomonitoring tool

Environment

**JOANNE E. LITTLEFAIR**
DEPARTMENT OF BIOLOGY, MCGILL UNIVERSITY, MONTRÉAL, CANADA

**JEAN CARREAU**
WSP CANADA

**MARK WEBB**
WSP UK

**MEIT ÖBERG**
WSP SWEDEN

**MELANIA CRISTESCU**
DEPARTMENT OF BIOLOGY, MCGILL UNIVERSITY, MONTRÉAL, CANADA

## Introduction

Aquatic systems are under mounting pressure from anthropogenic impacts, with associated rapid biodiversity loss. The WWF 2016 Living Planet report described biodiversity losses of 81% for global populations of freshwater organisms between 1970 and 2012, based on data from 3324 monitored populations from around the globe. Robust and up-to-date estimates of biodiversity are needed for researchers to effectively describe and monitor these losses, and propose solutions for conservation. Biomonitoring activities, which are often enforced in law, incur vast expense (e.g. Canada spends $500m each year on environmental monitoring, according to the 2011 Canadian Commissioner of the Environment and Sustainable Development).

The detection of environmental DNA (eDNA) has the potential to revolutionise biomonitoring. Animals leave behind traces of eDNA in the environment as they pass through habitats, or interact with other organisms. Researchers can collect water, soil or other substrates, filter out the DNA and sequence it with the latest powerful DNA sequencing technologies to identify species. The sequences are then matched against reference libraries of DNA to assign species identities.

eDNA is used in two principle ways: to look for a single species of interest and sequence one species at a time, or to sequence an entire ecological community from a mixed sample. When a single species is targeted, this technique is applicable to questions such as the detection of rare species, indicator species, or invasive species. For example, studies have focused on monitoring the presence of the great crested newt (a protected species in Europe), or detecting invasive bighead carp in the Great Lakes, USA [1,2].

When a whole community is targeted, primers are used to target an entire taxon, such as vertebrates, mammals, or particular groups of invertebrates. The information generated is similar to presence/absence data from a biodiversity survey, with no a priori knowledge of the organisms in the habitat. This technique has also been used to reconstruct the patterns of historical populations through analysis of eDNA in sediment cores. For example, a multinational academic team recently reconstructed historical patterns spanning the last 50,000 years of vegetation types and diets of large herbivores in the Arctic [3].

## What advantages do eDNA techniques present?

Existing monitoring techniques can be inefficient, destructive, taxonomically or functionally biased, and heavily reliant on professional taxonomic identification and therefore expensive. Robust estimates of entire community assemblages with morphological surveying are likely to require high sampling effort and multi-species approaches. The use of eDNA presents several advantages for monitoring in freshwater and marine habitats. With eDNA, versatile and sensitive multi-species approaches can be deployed, which generally have higher detection capabilities [4]. Researchers are additionally able to sample organisms and life stages which elude traditional techniques. For example, juvenile and gamete DNA are included in the detection of eDNA from water. Species can also be detected which, due to their cryptic ecology, may be very difficult to observe. The non-invasive aspect of sampling will also be attractive when target organisms are rare or juvenile life stages, which should be protected from destructive sampling.

Comparative studies have shown that high detection capacities of eDNA mean that it is time-efficient, particularly at the sample collecting stage. For example, it has been shown that 19 samples of water (250 ml each) distributed over a 2.2ha reservoir can plateau sampled species richness, which might otherwise be collected with several days of fishing [4]. Yet researchers should not forget that the subsequent laboratory and bioinformatics processing of samples comprises the major part of an eDNA workflow.

The cost efficiency of eDNA techniques in comparison to morphological surveys will to a great extent depend on the organism to be surveyed and the expense involved in deploying those techniques (Table 1). eDNA can entail high start-up costs, such as laboratory set-up and advanced expertise, although organisations without access to existing eDNA labs could collaborate or purchase expertise from academic or consulting research groups. The development phase of assays can equally be a time-consuming and somewhat costly process. However, once assays are developed, either de novo or by selecting primers from the literature, samples can often be processed in high-throughput workflows. Depending on the complexity of the communities to be analysed, traditional taxonomic identification can often be the rate-limiting step in biodiversity assessments.

*Table 1: The cost efficiency considerations of biomonitoring with eDNA and morphological surveys, with broad estimates of proportional cost for each area.*

| Cost efficiency considerations | eDNA | Morphological survey |
|---|---|---|
| Start-up costs | High. Dependent on the availability of molecular biology labs and high performance computing. Costs could be offset by collaborating with academic or consulting research groups. | System-specific. High for situations requiring remote monitoring, electro-fishing, or radiotelemetry. |
| Sampling | Inexpensive. Involves only the sampling of water or soil substrate. | Medium. Likely to involve more personnel-hours setting out traps and retrieving organisms from them. |
| Sample processing | Medium-high. Involves molecular biology consumables and potentially DNA sequencing costs. | Probably inexpensive. Preservatives and equipment used in sample processing are not costly. |
| Personnel | Medium. Dependent on molecular biology and bioinformatics expertise, which are relatively common amongst qualified scientists, although specialist training may be required. | Medium-high. Dependent on taxa-specific expertise, and the taxonomic resolution required. |

Lastly, the development of online resources means that eDNA workflows are auditable. Scientists are encouraged to upload raw data into repositories, which in the context of eDNA would involve information on sampling strategies as well as the raw sequencing data. Increasingly, scientists are maintaining open online lab books, and publishing bioinformatics scripts and tools which have been used to generate the results. Reproducibility of data can be verified, and used in subsequent studies of wider geographic regions or meta-analyses.

## The detection of species with eDNA: creation, transportation and decay

### Creation

eDNA is created by a combination of shedding skin cells, excretion, and saliva or decomposition of dead organisms. This DNA can be intra-cellular (still contained within cells, which makes up the bulk of detectable eDNA [5]), or extra-cellular (released from cells). Cells degrade due to environmental factors, and as they degrade, DNA is released and becomes extra-cellular.

### Transportation

eDNA is transported around the environment, especially in fast-flowing substrates such as water. This poses a problem for researchers wishing to understand presence/absence data in a flowing system, such as a river or lake chain. Understanding the answer to this question is the difference between detecting a fish present in the habitat, or one 1km upstream. Initial research suggests that eDNA can move between a few metres and possibly several kilometres under certain conditions. In lotic environments, it will certainly be difficult to tease apart the presence of species from the transport of eDNA from upstream habitats. Other vectors of eDNA transport could include predators, or man-made vessels such as boats.

### eDNA persistence

Mechanical forces, microbial action, heat, UV, enzymes, and acidity all accelerate the degradation of eDNA. To quantify degradation rates, laboratory studies have been conducted involving the removal of organisms from tanks or mesocosms, and subsequent measurement of species detection with eDNA. Studies in controlled aquatic environments, such as experimental ponds, suggest that eDNA remains in the water column for two weeks to two months. In the field, there is relatively little information about the persistence of eDNA once the species is extirpated. For example, eDNA seems to persist in sediments and soils for longer than in the water column, and indeed, sediment cores can be used to examine the historical presence of populations. The temporal longevity of eDNA in sediments is at present unknown.

## Methods

### Metabarcoding

The identification of species from eDNA generally relies on DNA barcoding. Barcodes are small regions of DNA often found in the mitochondria or chloroplast regions, which help with species identification. When multi-species detection is the goal of the study, DNA from the entire community is amplified using group-specific primers. Group-specific primers could include all invertebrates, vertebrates, fish or mammals, for example. These primers are short molecules which are designed to target a variable region of DNA flanked by two conserved regions which allow the primers to bind. The variable region must be targeted to contain sufficient variation to allow the identification of individual species. Once DNA is amplified, high-throughput sequencing is used to sequences millions of copies of DNA in parallel. Species are identified from a mixed sample by matching this DNA to reference libraries. The open source libraries are grown by campaigns to add reference barcodes for target taxa (e.g. PESTBOL - adding barcodes of plant pest species, FISHBOL - targeting fish species etc).

### Quantitative PCR

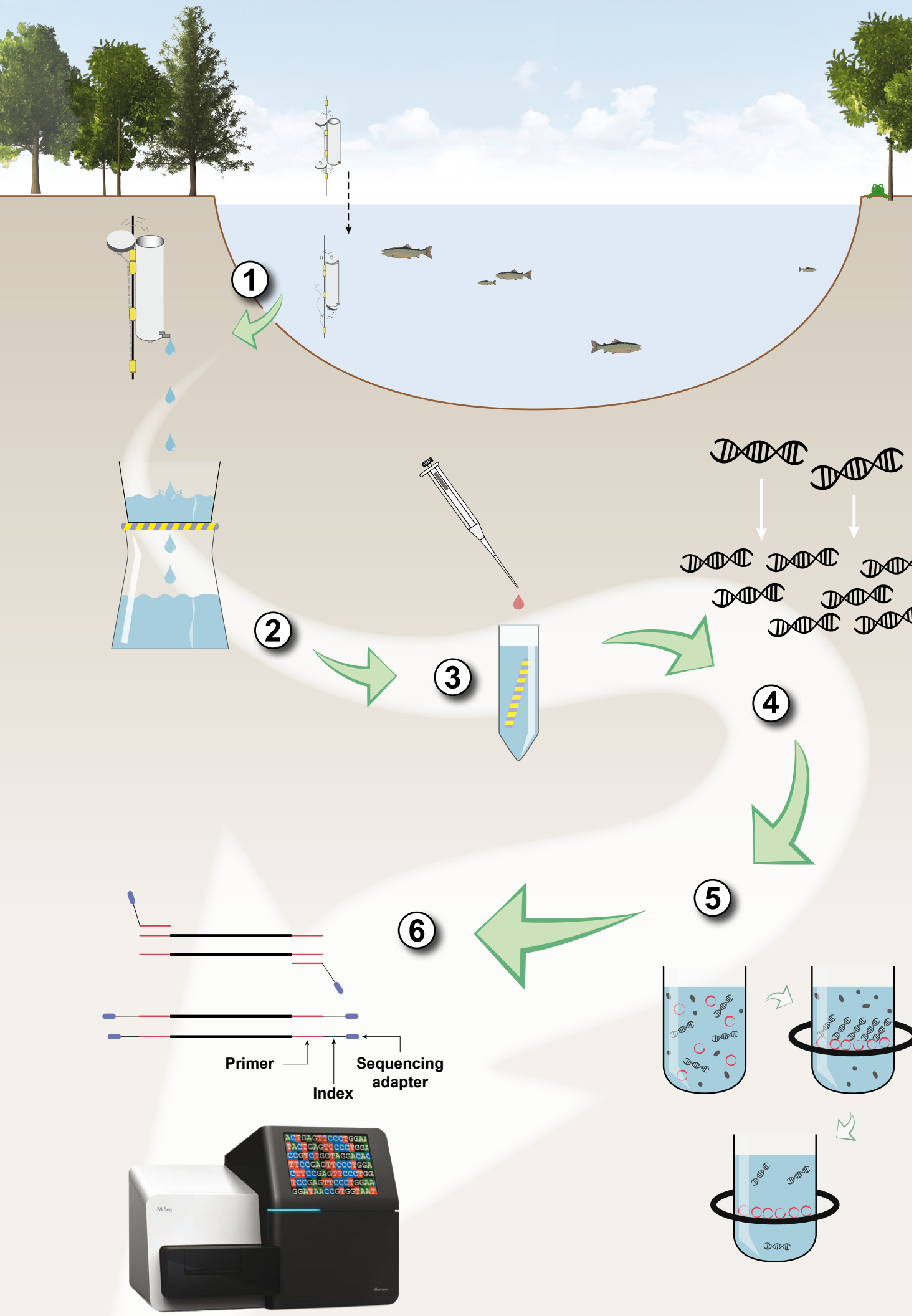
Quantiative PCR, otherwise known as real-time PCR or qPCR, is a quantitative technique which amplifies and detects DNA simultaneously. qPCR has many applications across molecular biology, although in the context of eDNA assays, it is used to detect a single pre-selected species. The detection of single species of interest relies on the use of primers which only amplify the DNA of that species, which must therefore be carefully evaluated for specificity against congeneric DNA.

Quantification takes place by attaching fluorescent probes to the DNA as it is amplified, and measuring the accumulation of a fluorescent probe during the exponential phase of the reaction. qPCR reactions are run on real-time PCR machines capable of detecting the fluorescent probe, but aside from this, the other components of the reaction are the same as regular PCR amplifications. The output from qPCR is the CT parameter, which is the number of cycles the fluorescence takes to exceed pre-defined thresholds, which will give estimates of relative abundance when compared with a calibrated system. In contrast, regular PCR is at best semi-quantitative, if the product is assessed by examining the brightness of bands on a gel.

qPCR is a cost-effective technique for detection, because it removes the sequencing stage from the bulk of the laboratory workflow. This technique is appropriate, for example, when searching for an invasive species, in particular when many samples have to be processed quickly and cost-effectively.

## CAN WE DETECT THE ABUNDANCE OF ORGANISMS IN A HABITAT USING eDNA?

*Estimating the abundance of organisms within a habitat is very important for ecological surveys and assessment of biodiversity. Although it has been attempted, linking molecular techniques such as metabarcoding and qPCR to population abundance is a contentious technique. In metabarcoding studies, researchers have attempted to link the number of sequences (read number, or copy number) to species biomass, but this often bears no relationship. Copy number may be influenced by dead or injured organisms, or by the presence of juveniles in the environment who may produce more eDNA per biomass because of their elevated metabolism during growth. Copy number of genes can also vary by an order of magnitude between species. More reliance can be placed on within-species relative abundances, estimated with probe-based PCR techniques such as qPCR, and digital droplet PCR, which detect only a single species (unlike metabarcoding). However, these techniques are likely to require system-specific calibration against relative abundance or biomass estimates such as CPUE or BPUE. It may therefore be more appropriate to categorise eDNA as a semi-quantitative technique, for example, distinguishing between abundant and rare species.*


1
2
3
4
5
6
Primer
Index
Sequencing adapter

## Methods to analyse eDNA: lab techniques for the metabarcoding of aquatic samples

1. ***Sample collection:*** The two most common methods for extracting eDNA from water samples are filtration and precipitation. Precipitation involves collecting a small amount of water (e.g. 15ml) and adding a salt solution and ethanol for preservation. Filtration involves suction of larger volumes through cellulose nitrate or glass fibre filter papers using syringes, hand-held pumps, or vacuum pumps. The volume of filtered substrate has varied amongst previous studies (from 15ml to 100L). The relationship between water volume and detection probability is as yet uncertain. One experiment attempting to quantify water volumes and detection probabilities found that no relationship existed for two species, but a positive relationship between water volume and detection probability did exist for one species [6]. As eDNA is known to be heterogeneously distributed throughout the environment, a wide spatial distribution of smaller samples is thought to be more representative of biodiversity than a single very large sample.
2. ***DNA extraction:*** A number of commercial kits have been used to extract DNA, including Qiagen Blood and Tissue kit, and the Powersoil and Powerwater kits from MoBio. CTAB-chloroform and phenol-chloroform extractions are also commonly used, and may present an advantage due to their high DNA yields and low cost.
3. ***PCR (polymerase chain reaction):*** PCR is required to amplify targeted eDNA regions to many millions of copies (amplicons) using primers (see Primer choice text box). Many millions of copies are needed so that the DNA sequencer can detect and read the DNA.
4. ***Purification:*** PCR amplicons are purified to remove small, non-target fragments.
5. ***Indexing:*** Amplicons are indexed, involving a second PCR reaction which incorporates short sample-specific sequences into the primer design. This enables all the amplicons to be pooled and sequenced together, thus reducing sequencing costs. Samples can later be separated during subsequent data processing, using these unique tags. When these samples have been tagged and pooled in equimolar concentrations, this is known as "library preparation".
6. ***DNA Sequencing:*** This is the process of reading the order of nucleotides within the DNA molecule. Metabarcoding sequencing takes place on high-throughput sequencing platforms (e.g. IonTorrent, Illumina, PacBio) - sometimes also known as next-generation sequencing technologies. Illumina platforms generate about 25 million reads per run - this is divided by the number of samples present. Therefore, increasing the number of samples in the run will decrease sequencing coverage, meaning rare sequences could be missed.

## WHICH PRIMERS SHOULD I USE?

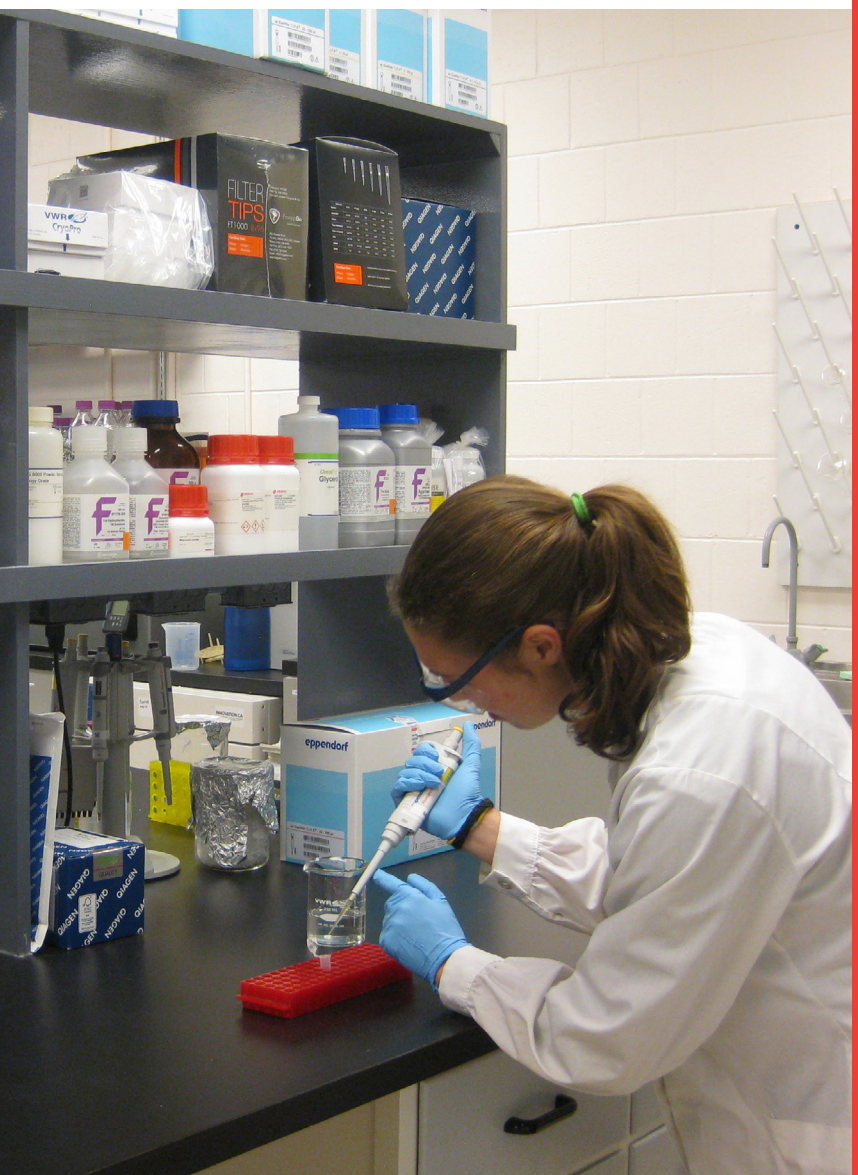
*Primer selection is an important part of project development for eDNA studies, because they ultimately affect the range of species detections possible. A number of primer sets are already available, but as many eDNA studies focus on detecting novel taxa, researchers may often supplement this with newly designed primers. Software is available to help choose primer regions and confirm taxonomic coverage. Commonly targeted regions should be selected for primer design, because of the reliance on reference libraries for taxonomic assignment, e.g. COI for vertebrates, 18S for other eukaryotes, 16S for bacteria. Historically, small regions have been targeted because of the degraded nature of eDNA, although generally longer fragments provide more taxonomic coverage. Careful selection of a primer set is important to minimise taxonomic bias by amplifying the DNA of all species present in the sample equally, while maximising taxonomic coverage. Generally, extensive in silico and in vivo validations are required to develop a primer set with these goals in mind.*

### Methods to analyse eDNA: bioinformatics

The output of high-throughput sequencing comprises many millions of raw DNA sequence reads. These data needs to be cleaned, filtered, assigned back to the correct samples, and assigned a taxonomic identity. Analytical steps will vary with the type of data generated (which could be dependent on sequencing platform, or the target taxa). A typical workflow is presented below:

1. ***Demultiplexing:*** This uses the sequencing indices attached to the reads to group the sequences according to their originating samples.
2. ***Trimming:*** Primers, sample indices, and sequencing adaptors should be removed, leaving only the species- or genera-specific sequence.
3. ***Length filtering:*** PCR primers amplify fragments of a specific length. Fragments that are much longer or shorter than the specified length could be sequencing errors.
4. ***Removal of chimeras:*** Chimeras are single sequences that originate from multiple parent sources and are created as artifacts in the PCR process. They are usually considered to be contaminants in eDNA workflows because they are misinterpreted as novel organisms and artificially inflate diversity. Removal can be performed computationally with several available software packages including QIIME.
5. ***Clustering:*** An optional step which involves the clustering of similar reads into Molecular Operational Taxonomic Units (MOTU), which correspond to a taxonomic unit. This also reduces processing time at the taxonomic assignment stage (see below). Clustering can be performed at different thresholds (e.g. 97% similarity) and by different clustering algorithms.
6. ***Taxonomic assignment:*** Sequences or MOTU are interrogated against a reference library, and taxonomic identities are assigned. Examples of large open reference libraries include SILVA (18S RNA), BOLD (eukaryote COI DNA), and UNITE (fungal ITS region).

Practitioners might use a number of bioinformatics programmes to perform these tasks, which require a moderate to high knowledge of command-line computing, although some packages are available for the analysis of metabarcoding data such as QIIME and OBITOOLS. Changes in either lab or bioinformatics pipelines can produce considerable variation in the eventual results. One challenge is to standardise the lab and bioinformatics workflows amongst research groups, to facilitate comparison between past and future studies and biomonitoring projects.




## REPLICATION

*The process of sample collection, DNA extraction, and PCR amplification is recognised to be subsampling true biodiversity. Therefore, replicated samples are needed at each stage (for example, multiple PCR replicates for each DNA extraction, which is recognised to be a stochastic process), where the amount of replication required is dependent on the study design. The number of replicates in eDNA studies at each stage varies from one to twelve, with many studies replicating in triplicate [6]. The fate of sample replicates depends on the study hypothesis and detection probability of the species. For example, if the aim of the study is to detect rare or invasive species, sequences might be included even if they just appear in only one replicate. Other studies might remove them. For example, if a study is working with highly degraded DNA, many cycles of amplification are required, thus inflating the chance of false positive detection. These researchers might choose to sequence PCR replicates separately, and confirm species presence in more than one PCR replicate, considering it dubious if the species does not appear in more than one replicate.*

## CONTAMINATION

*The process of capturing and extracting eDNA is sensitive to contamination in the field and lab. A number of processes are recommended to mitigate the risk of contamination:*

- *Preparation of sterile field kits for sample capture*
- *Thorough decontamination of field equipment between sites*
- *Extraction of field negative controls and lab negative controls*
- *Separation of pre- and post-PCR lab workflows, including separate freezers, workspaces, lab equipment and consumables. This prevents high-copy PCR products from contaminating low-copy samples and DNA extracts.*
- *Thorough and regular decontamination of lab equipment and workspaces, using UV hoods, bleach, and DNA removal agents.*
- *Advanced labs are equipped with whole-room UV, positive air flow, and are specialised to only process eDNA and aDNA (ancient DNA) work.*

### CASE STUDY: Using eDNA to survey and monitor the great crested newt (Triturus cristatus) in the UK.

The great crested newt is a legally protected amphibian species in the UK (protected by the Wildlife and Countryside Act 1981 (as amended), the Conservation of Habitats and Species Regulations 2010 (as amended), and the EC Habitats Directive. Conventional surveys for great crested newts are relatively reliable but strictly prescribed and labour intensive. They are carried out in line with industry guidance, the Great Crested Newt Mitigation Guidelines [7]. For a presence/absence survey, a minimum of three survey methods (from egg searching, netting, bottle trapping, torchlight surveys and searching terrestrial refugia) must be used on four survey occasions. There are strict seasonal constraints, with all surveys to be carried out between late March and early June (when the minimum overnight temperatures are greater than or equal to 5°C) and two of the four surveys taking place between mid-April and mid-May. Owing to the great crested newts' protected status, surveys must comply with this guidance as their presence (or absence) is a material consideration during the development consenting process.

In 2014, Natural England authorised the use of eDNA surveys for great crested newts as an alternative to conventional surveys. This followed UK Government funded research (through Defra, the Department for Environment, Food and Rural Affairs) carried out as a collaborative effort between the Freshwater Habitats Trust (UK), Natural England, Spygen (a DNA barcoding consultancy), the University of Kent (UK), and the Amphibian and Reptile Conservation [8]. The research team went on to develop a national citizen science-based monitoring programme for the great crested newt in the UK [2].

The scientists compared detection of newts with eDNA to conventional survey techniques. For the eDNA work, approximately 20 water samples were taken at each pond, which were based over two sites in southern England and north-east Wales. The water was combined, centrifuged, and the DNA pellet was extracted with a protocol modified from the Qiagen Blood and Tissue kit. qPCR was performed with

previously validated primers, which amplify an 81bp fragment of the cytochrome b subunit. Lab validation of these primers discovered that they amplify great crested newt DNA, but also with some mismatches to other fish and amphibians. Although this cross-amplification can be problematic, in this case, the researchers were able to use the predicted geographic distributions of the other species to exclude them from the analysis. The eDNA techniques detected newts in 139 out of 140 visits to 35 ponds, which great crested newts were known to inhabit (Table 2).

*Table 2: Success rate of detecting newts using either molecular methods or traditional surveying techniques, of 140 visits to 35 UK ponds where newts are known to be present.*

| Method of newt detection | Success rate of method |
|---|---|
| eDNA | 99.3% |
| Torch counts | 75% |
| Bottle trapping | 76% |
| Egg searches | 44% |

86 volunteer citizen scientists subsequently collected water from 239 ponds, some of which were known to be inhabited by the great crested newt. Ponds were distributed over the entire great crested newt range in England, Scotland and Wales. Citizen scientists were provided with sterile sample packs, but otherwise received no formal training. They collected 30ml samples from 20 locations within each pond and pooled this to make a 600ml sample. The eDNA was extracted and amplified in the same way as the formal comparison survey, and great crested newts were found to be present at 218 sites. Because the Limit of Quantification could not be reached for any qPCR, a rough attempt at estimating newt abundance was made by estimating the proportion of positive qPCRs out of 12.

Since the spring of 2014, WSP ecologists in the United Kingdom have been carrying out eDNA surveys for great crested newts to establish presence or absence. Where present, conventional surveys (based on a standard six-visit survey protocol), are implemented to establish population size class, on a scale of low, medium and high, to inform the statutory mitigation licensing process. Utilising eDNA is proving quicker and more cost effective for their clients, and of clear benefit to project programmes, given the potentially significant consequences of great crested newts for developments.

Furthermore, eDNA has been used by WSP ecologists in Sweden to survey six species of amphibians (table 3) in four projects during the spring of 2017. These projects have ranged from small-scale spatial planning to large-scale infrastructure development. The eDNA surveys provided basis for suggested mitigation measures, environmental impact assessment and potential need for further surveys. The use of eDNA to survey amphibians provided a more time efficient method as compared to conventional surveys, especially when performing preliminary screening for presence or absence of protected species at an early stage of development.

*Table 3: Amphibian species list survey used by WSP colleagues in Sweden*

| Species |
|---|
| Great crested newt (Triturus cristatus) |
| Smooth newt (Lissotriton vulgaris) |
| Common frog (Rana temporaria) |
| Moor frog (Rana arvalis) |
| Agile frog (Rana dalmatina) |
| Common toad (Bufo bufo) |

## FUTURE RESEARCH: eDNA to assess fish biodiversity in anthropogenically impacted habitats - a collaboration between McGill University and WSP

Aquatic ecosystems are economically and socially critical to Canada, supporting essential services such as clean drinking water and commercial fisheries as well as recreational activities. More than 60% of Canada's GDP is directly dependent on water. However, the anthropogenic pressures currently applied to aquatic habitats can generate extreme environmental conditions (high loads of nutrients, metals, increased acidity) that can alter the supported communities and compromise the function of these ecosystems. Traditional methods for sampling fish communities are very costly, may have a significant share of inaccuracy and require specialized taxonomic expertise. WSP is frequently called upon to assess the ecosystem integrity of aquatic ecosystems and mitigate impacts following industrial development.

Environmental DNA (eDNA) is being rapidly developed as a biomonitoring technique, due to its high detection capacity, rich taxonomic coverage, and non-invasive nature. Yet recent studies have highlighted the need for field, lab and bioinformatics validation of eDNA, before we know whether it is a true reflection of local species richness. The collaboration between scientists based at McGill University and WSP will develop and validate genomic tools that will allow accurate and rapid monitoring of biodiversity. Results from traditional sampling techniques will be compared with biodiversity estimates based on refined metabarcoding approaches to describe fish diversity within an array of impacted aquatic habitats. Development will focus on two primary areas:

1. Primer development
2. Transport of eDNA in flowing systems

Fish metabarcoding primers will be developed and validated in the lab. Multiple markers will be used for multi-species detection. Mock communities (artificially created mixtures of DNA extracts where the species composition is known) will also be assembled from fish tissue to complete lab validation. A primer cocktail will be optimised for the detection of freshwater fish in Québec ecosystems.

In flowing water systems (e.g. rivers, lake chains), local detection of species with eDNA is confounded by the transport of eDNA from upstream. Disentangling these effects (regional versus local diversity) in natural field conditions is reliant on the presence of viable populations which are restricted to one defined area, and monitoring the movement of eDNA released from them downstream. We therefore propose to study this question in a unique ecosystem of lake chains at the Experimental Lakes Area (ELA), Ontario, Canada. This is a world-leading aquatic ecology facility, where species distributions within the lakes are extremely well characterised, because of the decades-long intensive research program. Within the lake chains identified for study, some fish and zooplankton species are present in certain lakes but not others. We can therefore be very confident of species presence/absence, and that eDNA detections where species are absent are due to the transport of eDNA from other lakes. In addition, other factors which are known to influence the persistence of eDNA, such as water chemistry and hydrology, are very well characterised. Species composition and population abundance in the reference lakes are known in great detail.

Many questions still remain about the persistence and transport of eDNA in the environment, and whether these factors influence our understanding of patterns of species richness. Our study moves beyond experimental laboratory systems to a natural field setting to further our understanding of the ecology of eDNA. The results from this study will pave the way for considering eDNA as a measurement of local species richness, or alternatively, of regional-scale richness.

## Limitations of eDNA approaches

### What can eDNA tell us about species presence and absence?

The transport of eDNA and time delay in decay means that fine scale spatial and temporal inference of species presence and absence is challenging. In lotic systems, the continuous flow of water may mean that only watershed or landscape scale inference of species presence is possible. eDNA is released by both dead and live organisms, and therefore may not indicate the presence of a viable population. Other molecules, such as eRNA are only released by live organisms that are actively transcribing their DNA and could help researchers detect the immediate presence of a species. Although these molecules have been suggested for fine scale spatial and temporal inference of presence, the development of these assays is in its very infancy. Further validation will also have to be conducted in terms of how they match to the location of viable populations of organisms in the field. Currently, neither of these approaches can give us information about the population size or structure, population health or sex. This information will be critical for some applications, such as fisheries management, where traditional techniques are likely to provide more relevant information.

## False positives and negatives

*Table 3: Sources of false positives and negatives in eDNA work, and possible mitigation strategies*

| False positives | Mitigation strategies | False negatives | Mitigation strategies |
|---|---|---|---|
| Field contamination produces a false positive | 1. Careful sterilisation of all field equipment between sampling points, use of disposable consumables where possible.<br>2. Use of negative controls in the field protocol<br>3. Collection of samples at field sites where the species is known to be absent | Species is rare and therefore not detected | 1. High sequencing depth needed<br>2. Spatial, temporal and experimental replicates (e.g. increased PCR replicates) are needed to confirm species presence. |
| Lab contamination produces a false positive | 1. Rigorous maintenance of a contaminant-free laboratory environment<br>2. Use of separate pre- and post-PCR rooms, and a uni-directional workflow<br>3. Use of negative controls at each stage in the laboratory protocol | Primer bias excludes certain taxa | 1. Primer selection and validation against the relevant taxa. If a cocktail is used, the primers in question must be tested together |
| Sequencing error produces a false positive | 1. Bioinformatic procedures to remove singleton reads (sequences where there is only one copy)<br>2. Bioinformatic procedures to remove chimeras. | Overly stringent bioinformatics protocols excludes a species where it actually exists | 1. Validation of bioinformatics procedures against a mock community (an artificially created mixture of DNA extracts where the species composition is known) |

## Comparisons within and between eDNA studies

A major challenge for eDNA researchers is to create standardised lab and bioinformatics protocols in order to make results comparable within and between studies. One study might take place over a number of years, for example, an ongoing biodiversity assessment. Alternatively, researchers might wish to compare between a number of biodiversity assessments over a wide spatial distribution. Both require comparative data.

Comparisons are made more challenging because metabarcoding assessment for biodiversity studies is reliant on high-throughput sequencing technologies, where new reagents and platforms are regularly developed and released. Like the development of any new technology, bioinformatics tools are also constantly being developed and released. Many of the guidelines for eDNA studies will rely on careful validation against mock communities and field data at every step of the protocol.

## DNAQUANET - AN EU ACTION ON GENETIC TOOLS FOR BIOASSESSMENT

*In Europe, the EU Water Framework Directive makes the preservation and restoration of aquatic systems legally binding. The ecological status of water bodies are assessed through monitoring biotic indicators and comparing to a reference water body. DNAqua-Net is an EU Action called "Developing new genetic tools for bioassessment of aquatic ecosystems in Europe". It aims to develop novel genomic tools for biodiversity assessment with a unified approach across field and lab protocols, sequencing platforms, and indices. DNAquanet unites researchers from many European countries alongside international partner countries to form five working groups tackling: 1. DNA barcode reference databases, 2. Biotic indices and metrics, 3. Field and lab protocols, 4. Data analysis and storage, 5. Implementation strategies and legal issues. DNAquanet expects that molecular methods will replace traditional lab and field methods for surveying biodiversity in the next 10-15 years. Although national species inventories are in progress using DNA barcodes, it is hoped that a European-wide initiative will unify this procedure. The US Environmental Protection Agency has implemented a similar roadmap for the inclusion of molecular based methods into biodiversity assessment, which has been ongoing since 2011.*

## Other molecular methods for species detection

### Bulk samples

A metabarcoding approach can be applied to the original source material, as well as extracellular eDNA. Bulk samples, such as collections of floral pollen, zooplankton tows, and malaise trap content, can be extracted directly, and subject to the same metabarcoding procedures as outlined in this report. This type of biodiversity assessment might suit sampling strategies which expect to collect large numbers of species under conditions which make it difficult to apply large-scale morphological identification.

### Sediment cores

Sediment cores from lakes and glaciers can be used to detect the historical presence of populations using metabarcoding. This holds particular interest for those studying long term climate dynamics, glacial cycles, and reconstructing past environments. This delicate work with highly degraded DNA necessitates the highest standards of contamination prevention, and is typically conducted in specialised ancient DNA (aDNA) labs. Researchers are faced with unique challenges when working with sediment cores, including the presence of chemicals in cores which inhibit DNA extraction and PCR. Due to the degraded nature of the DNA, the best barcoding approach may be to target several barcoding markers with general primers, in order to obtain the maximum amount of taxonomic coverage.

### Molecular dietary analysis

Metabarcoding techniques have been refined for studies of dietary analysis, providing a very powerful means to disentangle otherwise cryptic trophic interactions. The same techniques of DNA extraction, amplification and sequencing can be applied to faecal matter, oral regurgitates, or gut contents. This is extremely helpful when an organism's feeding behaviour is secretive or otherwise difficult to observe, or when the diet is unidentifiable through morphological means; for example, liquid feeders. Studies of these interactions on a large scale can help researchers to build up highly resolved ecological networks, something which is only becoming possible with molecular techniques.

## Further reading

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

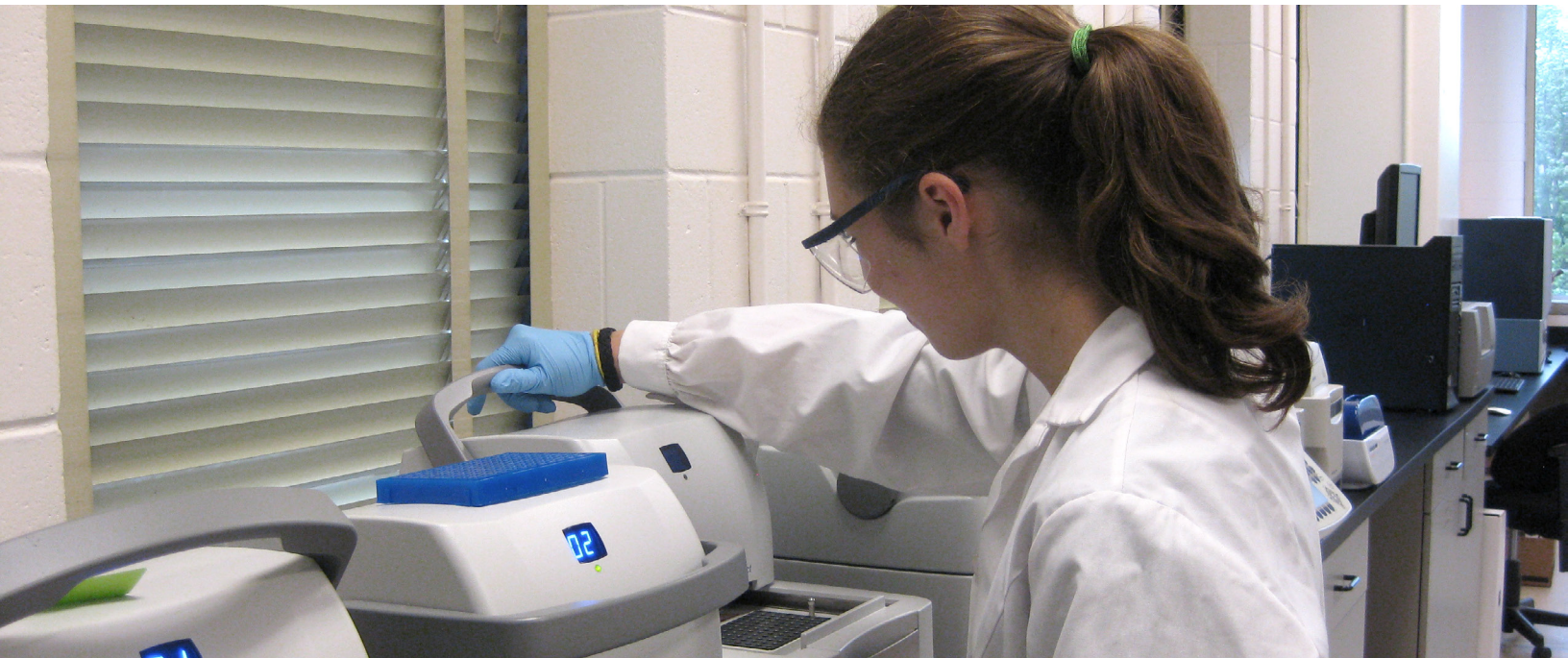
Cristescu, M.E. 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends Ecol. Evol. 29(10): 566–571. doi:10.1016/j.tree.2014.08.001.

Goldberg, C.S., Turner, C.R., Deiner, K., Klymus, K.E., Thomsen, P.F., Murphy, M.A., Spear, S.F., McKee, A., Oyler-McCance, S.J., Cornman, R.S., Laramie, M.B., Mahon, A.R., Lance, R.F., Pilliod, D.S., Strickler, K.M., Waits, L.P., Fremier, A.K., Takahara, T., Herder, J.E., and Taberlet, P. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods Ecol. Evol.: 1299–1307. doi:10.1111/2041-210X.12595.

Littlefair, J.E., and Clare, E.L. 2016. Barcoding the food chain: Sanger to high-throughput sequencing. Genome 59(11): 946–958. doi:10.1139/gen-2016-0028.

Lodge, D.M., Turner, C.R., Jerde, C.L., Barnes, M.A., Chadderton, L., Egan, S.P., Feder, J.L., Mahon, A.R., and Pfrender, M.E. 2012. Conservation in a cup of water: Estimating biodiversity and population abundance from environmental DNA. Mol. Ecol. 21(11): 2555–2558. doi:10.1111/j.1365-294X.2012.05600.x.

Shaw, J.L.A., Weyrich, L., and Cooper, A. 2016. Using environmental (e)DNA sequencing for aquatic biodiversity surveys: a beginner's guide. Mar. Freshw. Res. (May). doi:10.1071/MF15361.



## References

1. Klymus KE, Richter CA, Chapman DC, Paukert C. Quantification of eDNA shedding rates from invasive bighead carp Hypophthalmichthys nobilis and silver carp Hypophthalmichthys molitrix. Biol Conserv. Elsevier Ltd; 2015;183: 77–84. doi:10.1016/j.biocon.2014.11.020
2. Biggs J, Ewald N, Valentini A, Gaboriaud C, Dejean T, Griffiths RA, et al. Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (Triturus cristatus). Biol Conserv. 2015;183: 19–28. doi:10.1016/j.biocon.2014.11.029
3. Willerslev E, Davison J, Moora M, Zobel M, Coissac E, Edwards ME, et al. Fifty thousand years of Arctic vegetation and megafaunal diet. Nature. 2014;506: 47–51. doi:10.1038/nature12921
4. Evans NT, Li Y, Renshaw MA, Olds BP, Deiner K, Turner CR, et al. Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. Can J Fish Aquat Sci. 2017; 1–42. doi:10.1111/gcb.13051
5. Sassoubre LM, Yamahara KM, Gardner LD, Block BA, Boehm AB. Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish. Environ Sci Technol. 2016;50: 10456–10464. doi:10.1021/acs.est.6b03114
6. Machler E, Deiner K, Spahn F, Altermatt F. Fishing in the Water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. Environ Sci Technol. 2016;50: 305–312. doi:10.1021/acs.est.5b04188
7. English Nature. Great Crested Newt Mitigation Guidelines. English Nature, Peterborough UK. 2001
8. Biggs J, Ewald N, Valentini A, Gaboriaud C, Griffiths RA, Foster J, Wilkinson J, Arnett A, Williams P and Dunn F. Analytical and methodological development for improved surveillance of the Great Crested Newt. Appendix 5. Technical advice note for field and laboratory sampling of great crested newt (Triturus cristatus) environmental DNA. Freshwater Habitats Trust, Oxford. 2014

WSP™ is one of the world's leading professional services consulting firms. We are dedicated to our local communities and propelled by international brainpower. We are technical experts and strategic advisors including engineers, technicians, scientists, project managers, planners, surveyors and environmental specialists, as well as other design and program management professionals. We design and deliver lasting solutions in the Buildings, Transportation, Infrastructure, Oil & Gas, Environment, Geomatics, Mining, Power and Industrial sectors as well as project delivery and strategic consulting services. With over 7,500 talented people across Canada and 36,000 globally, we engineer projects that will help societies grow for generations to come.

**wsp.com**
**environnement-canada@wsp.com**