Title:

Long-read sequencing for biodiversity analyses - a comprehensive guide

Authors:

Iliana Bista[1,2 *], Alexandra Lino[1]

1: Senckenberg Research Institute and Natural History Museum, 60325, Frankfurt am Main, Germany
2: Wellcome Sanger Institute, Tree of Life, Wellcome Genome Campus, CB10 1SA, UK

*****Corresponding author: Iliana Bista (ilianabista@gmail.com)

Published version DOI: https://doi.org/10.1111/2041-210x.70250

Abstract

1. DNA-based monitoring of biodiversity has revolutionised our ability to describe communities and rapidly assess anthropogenic impacts on biodiversity. Currently established molecular methods for biomonitoring rely heavily on classic metabarcoding utilising short reads, mostly through Illumina data. However, increasingly more studies use long-read sequencing technologies, such as Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio), for analyses of environmental DNA (eDNA) and DNA barcoding. These long-read sequencing approaches can be advantageous over existing methods by providing increased information content and opening new avenues for understanding biodiversity at a larger scale.
2. In this review we provide an overview of studies to date using long-read sequencing platforms for biodiversity analyses of eukaryotes, from eDNA and community metabarcoding, shotgun sequencing, and DNA barcoding. We also give detailed information on each step required for sample processing, data generation and analysis of long-read sequencing datasets for biodiversity applications. Even though the number of studies using long-read sequencing technologies is rapidly increasing, clear established guidelines for sample preparation or analysis of such datasets are lacking. Furthermore, a side by side review for both ONT and PacBio technologies across different applications has not been presented.
3. Long-read sequencing platforms provide possibilities for metabarcoding (PCR-based) of both short and long fragments, shotgun sequencing (PCR-free), or DNA barcoding. ONT platforms in particular can also allow real time data acquisition while the portability of the MinION instrument can support sequencing in the field. PacBio on the other hand can provide highly accurate reads and can be used to reliably address open questions in ecology and evolution for difficult to characterise taxa such as microeukaryotes. Streamlining the use of these technologies could enable sequencing of whole organelles or population level assessments from environmental samples, which would be a step-change for DNA-based monitoring.
4. Here we review existing applications, applied methodology, and future perspectives in the field. Overall we aim to facilitate the use of long-read sequencing technologies by the wider community, promote best practices for data generation and standardised data reporting that supports reproducibility, and to encourage open data policies and tools.

Keywords: biodiversity monitoring, DNA barcoding, eDNA, long-read sequencing, metabarcoding, ONT, PacBio, shotgun

1. Introduction

Traditional methods for biodiversity assessment rely on taxonomic identification of species, which require highly specialized taxonomic expertise while they are generally low in throughput, costly, and potentially invasive (Deiner, Bik, et al., 2017; Ruppert et al., 2019). A global lack of taxonomic expertise and the need to increase throughput has led to the development of DNA-based methods for biodiversity assessment over the last 10-15 years (Kelly et al., 2014; Thomsen et al., 2012). These methods enable novel, non-invasive, and cost-effective biomonitoring solutions which can be performed at a larger scale (Taberlet et al., 2012). This DNA revolution has been built on metabarcoding, which refers to high throughput sequencing of PCR amplified gene fragments of DNA derived from environmental (eDNA) or community samples (Alberdi et al., 2017; Deiner, Bik, et al., 2017). Metabarcoding is most commonly performed using short-read sequencing like Illumina, which has been the golden standard due to its high efficiency, low error rate, and cost effectiveness. Nevertheless, the maximum amplicon length allowed through Illumina sequencing has been ~550 bp (Deiner, Bik, et al., 2017), but note that the recently released MiSeq i100 supports 500 PE (1000 cycle) sequencing capability. Targeting short regions provides insufficient information content for taxon identification, especially for closely related species, and limits the number of available primer binding sites hence making primer selection more difficult (Egeter et al., 2022).

Increasingly more studies use long-read sequencing as an alternative to classic metabarcoding, through Oxford Nanopore Technologies (ONT) or Pacific Biosciences (PacBio) sequencers. These platforms perform single-molecule real-time sequencing (MacKenzie & Argyropoulos, 2023; Rhoads & Au, 2015), and provide options for longer insert sequences which are not limited to specific read lengths. Published literature of ONT or PacBio based biodiversity analyses, including applications and tools, has been rapidly growing (Table 1). Although long-read sequencing is not yet routinely used in biodiversity analysis, it is becoming more common and it has already been tested in different settings like aquatic ecosystems, soil or air samples, fungal profiling and more (Doorenspleet, Jansen, et al., 2025; Heeger et al., 2018; Maggini et al., 2024; Truelove et al., 2019; Urban et al., 2023).

Typical gene markers used for DNA barcoding are also commonly used for metabarcoding work thus enabling use of existing barcode refence databases (Andújar et al., 2018; Deiner, Bik, et al., 2017; Ruppert et al., 2019). These markers include the COI, as well as 12S, 16S, ITS and other genes to support detection of different taxonomic groups (Andújar et al., 2018; Miya et al., 2015; Tedersoo et al., 2022). Metabarcoding enables the simultaneous analysis of many species providing a broader understanding of the composition of whole communities (Ruppert et al., 2019). Using longer and more informative metabarcoding amplicons with ONT and PacBio for metabarcoding may enhance the taxonomic resolution by covering larger fragments or multiple marker genes. This could help improve phylogenetic resolution and characterisation of unknown species in complex communities (Doorenspleet, Jansen, et al., 2025).

Despite ONT suffering high error rates in the past, this issue has been mostly resolved through recent improvements in chemistry and bioinformatic tools, while sequencing costs have also improved. Furthermore, ONT offers a unique option, the portable MinION, which can be used in the field supporting real time data generation for biodiversity monitoring studies (Krehenwinkel, Pomerantz, & Prost, 2019; Maestri et al., 2019). The PacBio platform offers highly accurate reads with very low

error rates (Kucuk et al., 2023). Despite PacBio having been more popular for whole genome sequencing, it is also implemented for amplicon analysis of environmental samples and for DNA barcoding (Gueidan & Li, 2022; Jamy et al., 2022; Runnel et al., 2022).

Here we provide an overview of the current applications of long-read sequencing in biodiversity studies for both ONT and PacBio sequencing technologies. We discuss the areas of DNA barcoding of individual specimens, PCR-based (metabarcoding) and PCR-free (shotgun) sequencing applications from eDNA and community samples. We focus on eukaryotes, including a wide range of taxonomic diversity, though the coverage of different taxa varies due to the varying extent of current applications. To support practical method uptake by the community, we outline required steps for sample preparation and bioinformatic analysis as well as future perspectives for long-read based biodiversity analyses. Ultimately, we aim to facilitate both new and experienced users to fine-tune their work and to encourage further development in this field.

2. Nanopore sequencing

Nanopore sequencing was developed in the 1990s, but it became generally available with the release of the portable MinION instrument by Oxford Nanopore Technologies (ONT) (Branton et al., 2008; MacKenzie & Argyropoulos, 2023). ONT allows real-time single-molecule sequencing of short to ultra-long reads, ranging from 50 bp to >4 Mb. A nanoscale protein pore known as the “nanopore” (MacKenzie & Argyropoulos, 2023) is used to recognise individual nucleotides through the detection of electrical signals which are emitted as the molecule moves through the nanopore (Fig. 1A) (Wang et al., 2021). This movement, known as translocation, and its speed are controlled by a motor protein which is attached at the 5' end of the template strand (Deamer et al., 2016; Wang et al., 2021) (Fig. 1A). Initial ONT releases suffered very low accuracy levels (R6 60% accuracy level) (Rang et al., 2018), but newer releases have upgraded chemistries, different types of nanopores and motor proteins, and improved accuracy (R10 nanopores: 99% read accuracy, 400-450 bases/sec translocation rate)(Srivathsan et al., 2021; Wang et al. 2021).

Existing ONT instruments include the MinION, GridION, and PromethION (Fig. 1A). The MinION is the standard ONT flow cell (2,048 nanopores), which can be used individually with the portable MinION sequencing device. The GridION is a benchtop device suitable for medium-scale projects that can accommodate up to 5 parallel runs with MinION flow cells. The PromethION device runs with flow cells with 10,700 nanopores, which can be run in parallel on the same device in 24 or 48 modules, making this instrument suitable for larger projects. Finally, the Flongle is a smaller flow cell (126 nanopores), which can be adapted to use with either the MinION or the GridION, providing a faster and cost-efficient solution for a low number of samples. The amount of data generated by a flow cell during a run depends on the number of active pores, duration of the run, and translocation rate.

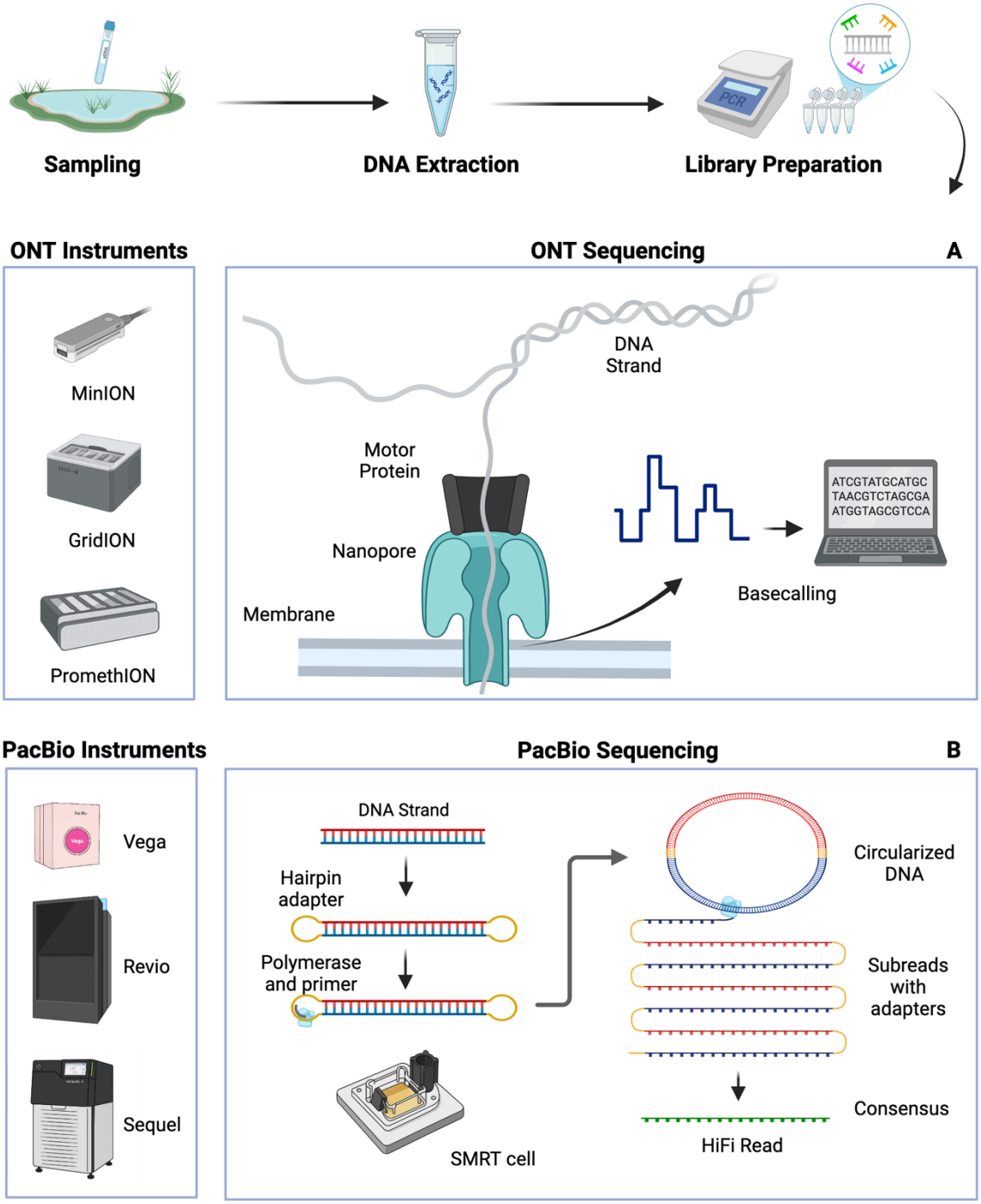


Figure 1

Sample processing workflow and sequencing with long-read technologies. A) ONT instruments and sequencing: movement of the DNA strand through the nanopore leading to data generation. B) PacBio instruments and sequencing: SMRT cell sequencing, showing the Circular Consensus Sequencing (CCS) process for generation of HiFi reads (Image generated with BioRender).

3. Applications of ONT for biodiversity analysis

3a. ONT DNA Barcoding

ONT sequencing for biodiversity analysis was first used for DNA barcoding of individual specimens covering different taxonomic groups such as terrestrial vertebrates (Menegon et al., 2017; Pomerantz et al., 2018), dipterans (Srivathsan et al., 2018), nematodes (Knot et al., 2020), marine invertebrates (Ip et al., 2023), insects, arachnids and plants (Krehenwinkel, Pomerantz, Henderson, et al., 2019).

Some important advantages of ONT DNA barcoding include its high multiplexing capacity and increased throughput. In practise, these advantages translate to reduced costs by lowering the cost per individual barcode (Hebert et al., 2025; Maestri et al., 2019; Pomerantz et al., 2018), which in turn can be used to dramatically accelerate the population of reference databases, which is a long-wanted resolution for the barcoding community (Hebert et al., 2018; Srivathsan & Meier, 2024). To date, this increased throughput has enabled huge leaps in the scale of barcode generation, as demonstrated for example with the generation of barcodes for >100,000 insect specimens in a single run (Hebert et al., 2025), or by country wide barcode generation efforts ("megabarcoding") (Hartop et al., 2024).

Comparisons of ONT vs Sanger generated barcodes demonstrated the same levels of accuracy (Koblmüller et al., 2024), while library preparation for ONT (vs Sanger and PacBio) was also shown to be less time consuming overall (Cuber et al., 2023). Srivathsan et al., (2019) sequenced ONT barcodes for more than 7,000 Dipteran specimens yielding 99.99% accuracy compared to the Illumina data. In this case though, comparison was only possible for a 313 bp subset of the full-length barcode that corresponded to the available length of the Illumina amplicon.

Considerable development has been made to enable ONT field-based applications using the portable MinION instrument. Seah et al., (2020) tested field-friendly DNA barcoding methods in different non-invasive wildlife samples, while Vasilita et al., (2024), evaluated express protocols for species-level sorting of insects, obtaining 250 barcodes in ~6 hours, with minimal laboratory equipment. Other such studies have been implemented in rainforest and national parks (Gygax et al., 2025). For example Pomerantz et al., (2018) tested the feasibility of a portable laboratory for real-time barcoding of reptiles in the Ecuadorian rainforest, whilst Maestri et al., (2019) implemented in situ DNA extraction and barcoding for frog and beetle species in the Temburong National Park, Borneo.

Long ONT barcode protocols have been developed aiming to improve downstream phylogenetic analysis. For example, sequencing of the full length ribosomal cluster (~4,000 bp) enabled the identification of arthropods with high phylogenetic resolution (Krehenwinkel, Pomerantz, Henderson, et al., 2019). Similarly, Gajski et al. (2024) sequenced large fractions of mtDNA and nuclear genes (COI, Cytb, 18S-28S) of arthropods with a multiplex PCR barcoding approach, which corresponded to more than 6,000 bp of data per specimen. This approach was considerably more efficient, producing much more sequencing data per specimen than combining individual Sanger-sequenced barcodes.

3b. ONT PCR-based eDNA and community sequencing

Classic metabarcoding analysis focuses on short amplicons, and even more so for eDNA samples where DNA fragmentation due to degradation occurs (Deiner, Bik, et al., 2017). Nevertheless, sequencing of longer than the typical short fragments and full fish mitogenomes with long-range PCR have been done from aquatic eDNA (Bista et al., 2017; Deiner, Renshaw, et al., 2017). Furthermore, longer fragments are highly abundant in community samples and can be sequenced with long-read platforms (Baloğlu et al., 2021).

Community sample sequencing with ONT has been performed in various settings. Toxqui Rodríguez et al. (2023) used artificial mixtures of seafood samples to evaluate the efficiency of ONT for detection of seafood fraud. Similarly, Voorhuijzen-Harink et al. (2019) and Ho et al. (2020) sequenced fish mixtures or highly processed seafood samples respectively for food authentication purposes. Additionally, Voorhuijzen-Harink et al. (2019) also demonstrated that the accuracy levels of ONT vs Illumina data from the fish mixtures were comparable. This was also shown for bulk zooplankton (Chang et al., 2024; Semmouri et al., 2021), macrobenthos (Doorenspleet, Mailli, et al., 2025), coral reefs (Ip et al., 2023), and bulk mosquito communities (Loh et al., 2024) which were sequenced with both platforms.

ONT sequencing from aquatic eDNA samples is commonly used for fish diversity monitoring and fisheries management (Table 1). In two cases where short and long amplicons were sequenced in parallel (170 & 2,000/2,400 bp), it was shown that higher detection rate was actually achieved with the shorter amplicons (Doorenspleet, Jansen, et al., 2025; Maggini et al., 2024). This could be due to lower availability of longer fragments in eDNA samples due to degradation (Doorenspleet, Jansen, et al., 2025). Nevertheless, longer fragments improved species assignment in closely related species, while overall detection rates were correlated to low species abundance and low genetic variance (Maggini et al., 2024). Finally, short amplicon ONT eDNA metabarcoding (170-200 bp) (Kasmi et al., 2024; Munian et al., 2024) was shown to outperform conventional methods of fish monitoring both in the overall diversity detected and in terms of cost efficiency. Aquatic eDNA metabarcoding demonstrated comparable results for detection of invasive bivalves (Egeter et al., 2022), or better results for vertebrate detection (Gygax et al., 2025), when Illumina and ONT results were compared. Furthermore, MinION was used for white shark detection in the open ocean, performing all steps of wet-lab and bioinformatic analysis on board ship within ~48 hours, thus highlighting ONT applicability for real-time evaluation of biodiversity (Truelove et al., 2019).

ONT has also been used for pollinator and pathogen monitoring in plants. Here, eDNA collected from flowers was used to assess the diversity of pollinators, which detected a large part of the pollinator diversity, though lower than Illumina and traditional methods (Harris et al., 2023). Furthermore, Theologidis et al. (2023) used ONT for detection of fungi and endophytes infecting young olive tree shoots as a potential phyto-diagnostic method. Finally, sequencing of both long and short ONT amplicons for the 18 rDNA region was used to characterise protist communities (Bludau et al., 2025; Gaonkar & Campbell, 2024).

3c. ONT PCR-free (shotgun) eDNA and community sequencing

PCR-free (shotgun) sequencing of eDNA or community samples offers an alternative as a non-targeted

biodiversity monitoring tool. Here sequencing is not limited to single gene markers but represents fragments of all nuclear and mitochondrial genomes present in the sample (Linard et al., 2015; Peel et al., 2019). This method bypasses potential primer bias, allowing a more universal view of the community composition of the sample (Bista et al., 2018). Few studies have applied this approach to eDNA samples so far. In Munian et al. (2024) PCR-free ONT sequencing was used for detection of freshwater fish from water samples, and similarly in Nousias et al. (2024) for assessment of riverine biodiversity. Furthermore, Koda et al. (2023) used shotgun sequencing of eDNA to detect the presence of bobcat DNA extracted from soil samples of pawprints, while Urban et al. (2023) performed genomic biomonitoring of the endangered kākāpō bird, by analysing soil samples collected from the areas that the birds frequented. Besides environmental samples, Shum et al. (2024) used shotgun ONT for fish and seafood traceability to confirm identity and geographic origin with a more rapid turnaround.

4d. ONT microbial analysis in biodiversity studies

Microbial community analysis is most commonly performed through 16S amplicon or shotgun (metagenomic) sequencing (Quince et al., 2017). Numerous studies have used ONT for microbial community sequencing focusing on clinical applications, for review see Kim et al. (2024), hence we briefly mention some cases where microbial analysis of environmental samples is used with biodiversity implications. For example, 16S amplicon ONT sequencing was used to assess the impacts of agriculture or aquaculture impacts in the environment (Stevens et al., 2023; Stoeck et al., 2024), or the potential health impact of dust microbiomes collected from air samples (Nygaard et al., 2020). Furthermore, an atlas of bacterial diversity in the River Cam, Cambridge was established with MinION, aiming to monitor water quality in relation to surrounding land use (Urban et al., 2021).

Additionally, shotgun metagenomic sequencing of microbial communities provides wider opportunities for community assessment, and the possibility of assembling whole microbial genomes (metagenome-assembled genomes, MAGs) (Quince et al., 2017). When shotgun sequencing is used, added value can be gained from the long reads, beyond the original aim of the project. For example, Liu et al. (2024) were able to detect a functional shift in microbial communities related to environment depth, while mainly aiming at investigating cichlid diversity in lake Masoko. In Clark et al. (2025) ONT metagenomics was used to assess levels of antimicrobial resistance (AMR) in nearshore sediments from Antarctica, while Reska et al. (2024) characterised the microbiome of air samples, confirming the potential use of ONT metagenomics in air eDNA studies.

221 **Table 1:** Overview of taxonomic target, gene markers, and amplicon length for published ONT and
222 PacBio studies discussed in this review.

Approach	Taxa targeted	Gene marker	Fragment size (bp)	Studies	Platform
DNA barcoding	Invertebrates Insects Amphibians Reptiles Birds Fungi Arthropoda	COI 16S Cytb rDNA	313-500	Pomerantz et al., (2018); Seah et al., (2020); Srivathsan et al., (2018, 2021, 2024); Ip et al., (2023); Vasilita et al., (2024), Hartop et al., (2024)	ONT
			600-10,000	Menegon et al., (2017); Srivathsan et al., (2018, 2019, 2021); Wurzbacher et al., (2018); Krehenwinkel, Pomerantz, Henderson, et al., (2019); Maestri et al., (2019); Cuber et al., (2023); Gajski et al., (2024); Koblmüller et al., (2024); Vasilita et al., (2024); Hebert et al., (2025)	ONT
	Arthropoda Invertebrates Fungi Nematodes	COI ITS LSU	650-10,000	Hebert et al., (2018); Wurzbacher et al., (2018); Runnel et al., (2022); Cuber et al., (2023)	PacBio
Metabarcoding	**Sample Type**	COI 12S 16S 18S Cytb ITS LSU	85-620	Voorhuijzen-Harink et al., (2019); Ho et al., (2020); Srivathsan et al., (2021); van der Reis et al., (2022); Chang et al., (2024); Gaonkar & Campbell (2024); van der Vorst et al., (2024); Loh et al, (2024); Bludau et al., (2025); Doorenspleet, Mailli, et al., (2025)	ONT
	Community		650-2,758	Ho et al., (2020); Semmouri et al., (2021); Baloğlu et al., (2021); Srivathsan et al., (2021); Toxqui Rodríguez et al., (2023), Dierickx et al., (2024), Gaonkar & Campbell (2024), Huggins et al., (2024), Bludau et al., (2025)	ONT
			585 - 6,300	Heeger et al., (2018); Orr et al., (2018); Tedersoo et al., (2018); Gueidan et al., (2019); Kolaříková et al., (2021); Jamy et al., (2020, 2022); Gueidan & Li, (2022)	PacBio
	eDNA		95-550	Truelove et al., (2019); Egeter et al., (2022); Harris et al., (2023); Theologidis et al., (2023); Munian et al., (2024); Kasmi et al., (2024); Maggini et al., (2024); Gygax et al., (2025); Doorenspleet, Jansen, et al., (2025)	ONT
			2,000-2,400	Maggini et al., (2024); Doorenspleet, Jansen, et al., (2025)	ONT
Shotgun	**eDNA**	n/a	n/a	Peel et al., (2019); Urban et al., (2023); Koda et al., (2023); Nousias et al., (2024); Doorenspleet, Mailli, et al., (2025)	ONT

4. Laboratory steps for ONT sequencing

Similar to other sequencing workflows, the standard steps needed for library preparation include DNA extraction, PCR amplification (if appropriate), labelling, and adapter tagging (Bruce et al., 2021). Library prep type for ONT can vary depending on the application such as using one or two PCR steps to amplify DNA before sequencing (Fig. 2), or attaching barcoded tags used for multiplexing many samples in a single run (Bohmann et al., 2022; Bruce et al., 2021).

4a. DNA extraction

The DNA extraction methods used vary based on the type of sample and the application. Ultimately what is important is that the extracted DNA meets the quality and quantity requirements needed for library preparation. eDNA samples are expected to contain lower concentration of DNA, with higher degradation compared to tissue extracted DNA used for DNA barcoding, while purification steps may be required for inhibitor removal (Bruce et al., 2021). Some of the most commonly used kits for extraction of environmental samples from soil or water are the DNeasy Blood and Tissue kit, the DNeasy PowerSoil, and the DNeasy PowerWater kit (Qiagen) (Supplementary Table 1). For field applications with MinION, the ease of DNA extraction should also be considered so that it can be performed in a non-lab setting. Such extraction methods which are fast and require minimal equipment, are the QuickExtract™ solution (Srivathsan et al., 2019) and the Chelex® 100 resin (Koblmüller et al., 2024; Seah et al., 2020). These methods might not always support the best DNA quality and quantity, so it is important to also assess the suitability of the DNA extract for downstream applications. If a large field collection is planned it is also advisable to perform trials before deploying in the field (Seah et al., 2020; Vasilita et al., 2024). Sample suitability tests include concentration and purity checks (Qubit or Nanodrop), and fragment length verification with agarose gels, Tapestation, or Bioanalyzer(Bruce et al., 2021).

4b. Library prep for PCR-based ONT sequencing

Following DNA extraction, the first step is amplification, labelling with different barcodes, and finally ligation of an ONT sequencing adapter to each barcoded amplicon (Fig. 2). For library preparation, three different strategies are commonly used in ONT amplicon studies (Figs 2. A, B, C).

1) Single-step PCR: All the samples are labelled directly during the first amplification with the use of primers which already include unique barcodes attached to the 5' end (barcode length 9 - 15 bp). This strategy has been used mostly in DNA barcoding studies (Gajski et al., 2024; Harris et al., 2023; Hebert et al., 2025; Ho et al., 2020; Koblmüller et al., 2024; Krehenwinkel, Pomerantz, Henderson, et al., 2019; Srivathsan et al., 2018, 2019, 2021; Vasilita et al., 2024), bacterial DNA from water samples (Urban et al., 2021) and zooplankton communities (Chang et al., 2024).

2) Two-step PCR with "tails": Target specific primers complemented with specific ONT tails are used for the first amplification. These tails will allow the binding of the different barcodes in a follow-up PCR reaction, commonly using only a low number of cycles. This strategy has been used in several DNA barcoding (Maestri et al., 2019; Pomerantz et al., 2018; Seah et al., 2020) and eDNA metabarcoding studies (Egeter et al., 2022; Huggins et al., 2024; Kasmi et al., 2024; Maestri et al., 2019; Pomerantz et al., 2018; Seah et al., 2020; Stevens et al., 2023; Theologidis et al., 2023).

3) Two-step PCR without “tails”: This strategy also uses the target specific primers but without additional tails in the first amplification. To label each sample an extra reaction is needed to modify the amplicons’ extremities and add barcode adapters (end-prep and barcode adapter ligation). Only after this reaction it is possible to perform the second PCR where the different barcodes are added. Fewer studies have used this strategy for DNA metabarcoding (Toxqui Rodríguez et al., 2023; van der Reis et al., 2022).

When extensive multiplexing is required, it is also possible to perform dual barcoding by using inner and outer barcodes. In this case the PCR Barcoding Expansion Kit is used (96 inner barcodes), followed by the Native Barcode kit (24 outer barcodes), which allows multiplexing up to 2,034 samples in the same flow cell (Supplementary Table 2). Performing clean-up steps between PCR reactions is required to remove non-target fragments, and primers from the previous reaction, and to ensure that only the intended amplicons are present in the final library (Fig. 2A). Usually magnetic beads (Ampure) are used for clean-up, which can discard byproducts based on size and according to beads-to-sample ratio.

In all strategies, after labelling, the samples are pooled equimolarly, and the final library is end-prepped, in a reaction that repairs and prepares the DNA ends for the sequencing adapter attachment (5' phosphorylated, 3' dA-tailed ends). Then a ligation reaction will attach the ONT sequencing adapters, consisting of an oligonucleotide and motor protein (Fig. 2B), which aids the translocation of the DNA strands through the nanopore (Fig. 1A). Finally, another clean-up step is used, either with Long Fragment Buffer (LFB) to enrich DNA fragments >3 kb, or with Short Fragment Buffer (SFB) to retain DNA fragments of all sizes. The finalised library is loaded and the run, which can be monitored in real time, can go on for up to 72 hours, or until the required number of reads has been achieved.

4c. Library prep for PCR-free ONT sequencing

For the PCR-free shotgun approach, the PCR amplification steps are omitted, and the end-prep and ONT sequencer adapter ligation can be done directly after DNA extraction (Koda et al., 2023). However, when multiplexing is used, a labelling step for the addition of individual barcodes needs to be included. This can be done with an ONT Rapid Barcoding kit that uses a transposase barcode complex to cleave gDNA and attach barcodes (Peel et al., 2019; Reska et al., 2024), or with a Ligation Sequencing kit with Native Barcodes that will attach barcodes to fragmented gDNA after an initial end-prep reaction (Munian et al., 2024; Reska et al., 2024).

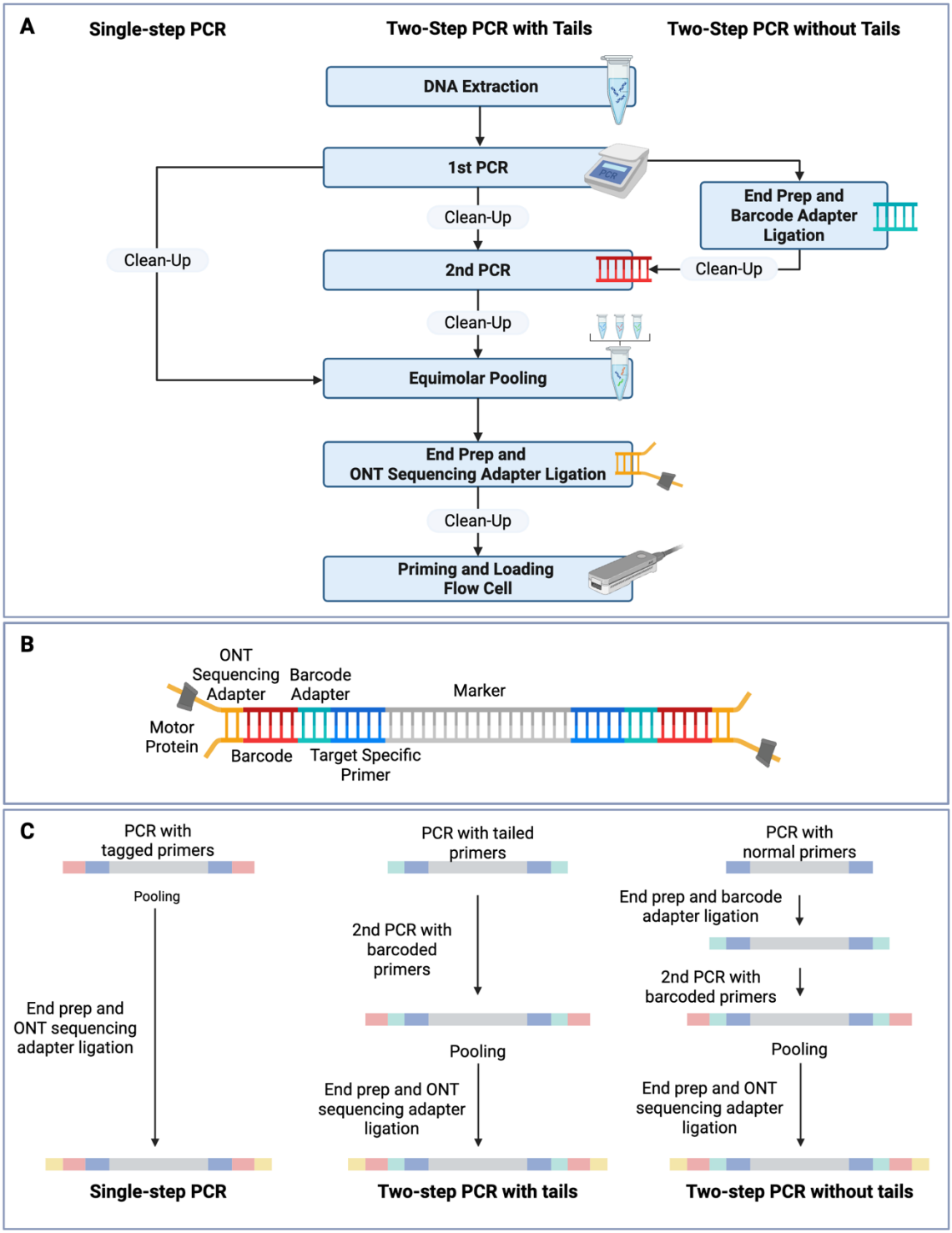


Figure 2

Strategies for library prep for ONT amplicon sequencing. A) library prep strategies: Single-step PCR, two-step PCR with tails, and two-Step PCR without tails B) structure of the amplicon with all types of adapters and barcodes, C) amplicon structure for each library preparation strategy. For PCR-free (shotgun) sequencing, the end prep and adapter ligation are performed without intermediate amplification steps (Image generated with BioRender).

298 4d. Sample requirements for ONT sequencing

299 Important considerations for maximising the success of an ONT run include the quality and quantity
300 of the sample, the type of flow cell and chemistry, and the respective kits. Keeping the occupancy of
301 nanopores high ensures optimal yield and is related to the amount of input DNA, as in number of
302 molecules available and their size. The loading amount of the final library depends both on the
303 fragment size, and the type of flow cell used. Shorter fragments will require larger quantities because
304 they are sequenced faster, while longer fragments keep the nanopores occupied for longer, hence the
305 required amount is lower. However, there is not a linear relationship between the amount of DNA
306 loaded into the flow cell and the sequencing yield, only a likely reduced output when these are
307 underloaded.

308 With the latest chemistry flow cells R10.4.1 (early 2025), loading approximately 35-50 fmol of good
309 quality library is recommended to achieve high pore occupancy (Doorenspleet, Mailli, et al., 2025;
310 Srivathsan et al., 2024). As these values also depend on the amount of DNA molecules and their size,
311 it is advised that specific ONT guidelines are considered based on the library fragment length (Table
312 2, Supplementary Table 2). Nevertheless, testing several input quantities of mock microbial
313 communities (1-1,000 ng), showed that lower DNA amounts than those recommended by ONT
314 guidelines could also be successful when applied to metagenomics (Simon et al., 2023). This aspect of
315 library prep for ONT metabarcoding has not been systematically explored to date, and comparison
316 between studies is challenging due to different flow cell types and fragment size used, but also due to
317 incomplete reporting. For future work it is important that information such as DNA input is disclosed
318 to enable further standardization and inter-study comparisons.

319 The most commonly used ONT kit for library prep for biodiversity studies is the Ligation Sequencing
320 Kit (SQK-LSK114), for which "barcoding" is performed beforehand either with ONT PCR Barcoding
321 Expansion kits (Doorenspleet, Jansen, et al., 2025; Egeter et al., 2022; Pomerantz et al., 2018) or with
322 customized/tagged primers (Hebert et al., 2025; Koblmüller et al., 2024; Krehenwinkel, Pomerantz,
323 Henderson, et al., 2019). For this kit 100-200 fmol of amplified DNA is recommended as the optimal
324 input (Doorenspleet, Jansen, et al., 2025; Srivathsan & Meier, 2024). The Native Barcoding and Rapid
325 Barcoding kits, which include the "barcoding" and don't require a PCR reaction, are less commonly
326 used. The Native Barcoding kits have a long library preparation time (~140 min) and are usually
327 recommended for gDNA samples (Munian et al., 2024), but can also be used with amplicons (Bludau
328 et al., 2025; Stoeck et al., 2024). The Rapid Barcoding kit has a shorter prep time (60min) because it
329 performs fragmentation and barcode attachment in the same step, though this might reduce yield
330 and read quality (Clark et al., 2025; Reska et al., 2024). See Table 2 for input requirements for the
331 Ligation Sequencing Kit, and Supplementary Table 2 for other kits.

Table 2

Input DNA amount for library prep and flow cell loading for ONT, according to fragment length and type of flow cell (Ligation Sequencing kit SQK-LSK114 kit). For other kits see Supplementary Table 2.

Fragment Length	Starting Input		Loading Amount	
	Flongle	MinION and PromethION	Flongle	MinION and PromethION
<1 kb	50 fmol	200 fmol	5-10 fmol	100 fmol
1-10 kb	50-100 fmol	100-200 fmol	5-10 fmol	35-50 fmol
>10 kb	500 ng	1 ug		300 ng

5. Bioinformatic analysis of ONT data

In contrast to standard length Illumina data, ONT (and PacBio) can generate both long and short fragments with varying error profiles (van der Vorst et al., 2024), hence bioinformatic tools need to be able to accommodate these differences. During ONT sequencing, reads are generated through the conversion of raw data collected in the form of electric signals into nucleotide sequences, called basecalling (Wang et al., 2021). After basecalling, quality filtering and processing steps are performed to obtain the final dataset according to each application (Fig. 3).

5a. DNA barcoding

In DNA barcoding each "barcode" corresponds to one species, and the objective is to generate a final consensus sequence for each specimen. After demultiplexing, the adapter and "barcode" sequences are trimmed, and the reads are filtered based on quality scores and size. A clustering step is usually performed at a lower percentage identity, since the variation in the reads from a single specimen is often due to sequencing error or contamination (Hebert et al., 2025; Maestri et al., 2019). Alternatively clustering is performed for each "barcode" only keeping the reads from the most abundant cluster or from clusters above a certain threshold (Maestri et al., 2019; Seah et al., 2020). Then a draft consensus sequence is generated for each sample which is polished, usually by mapping all the reads back to the draft consensus (Krehenwinkel, Pomerantz, Henderson, et al., 2019; Maestri et al., 2019; Pomerantz et al., 2018) or by re-clustering the sequences at a higher identity percentage and doing a new consensus calling (Hebert et al., 2025). The final consensus sequence is trimmed to obtain the final DNA barcode sequence and then aligned to reference sequences in databases (Hebert et al., 2025; Maestri et al., 2019). The order of the processing steps may differ between studies. The ONTbarcoder pipeline (Srivathsan et al., 2021, 2024) is one of the most commonly used in barcoding studies (Cuber et al., 2023; Harris et al., 2023; Koblmüller et al., 2024; Vasilita et al., 2024).

5b. eDNA and community sample metabarcoding

Initial preprocessing steps for eDNA or community metabarcoding include demultiplexing, adapter trimming, general quality control and size filtering, followed by clustering of Operational Taxonomic Units (OTU) or Amplicon Sequencing Variant (ASV) processing, and taxonomic assignment (for

363 metabarcoding tools see (Hakimzadeh et al., 2024) (Fig. 3). The exact order of these steps may differ
364 between studies. After demultiplexing, the barcode and adapter sequences are trimmed and reads
365 that do not have the correct primer sequence are discarded (Maggini et al., 2024). Quality assessment
366 and filtering of the reads is performed according to quality scores and expected fragment length,
367 removing low-quality reads and improving accuracy.

368 For OTUs, clustering is performed according to a similarity threshold, which could vary based on the
369 level of intraspecific diversity of the target groups (Alberdi et al., 2017), the ONT kit/chemistry,
370 basecalling algorithm, or fragment length (Doorenspleet, Jansen, et al., 2025). Due to ONT's error
371 profile, clusters with a low number of reads are often eliminated in order to reduce noise. However,
372 this step needs to be adjusted to each study depending on the target taxa and expected presence of
373 rare/low-abundance species. Subsequently, the reads in each cluster are aligned and a draft consensus
374 sequence is generated for each OTU, while polishing can be done to improve the accuracy of the
375 consensus sequences (Doorenspleet, Jansen, et al., 2025). Finally, an OTU table is built summarizing
376 OTU diversity and abundance.

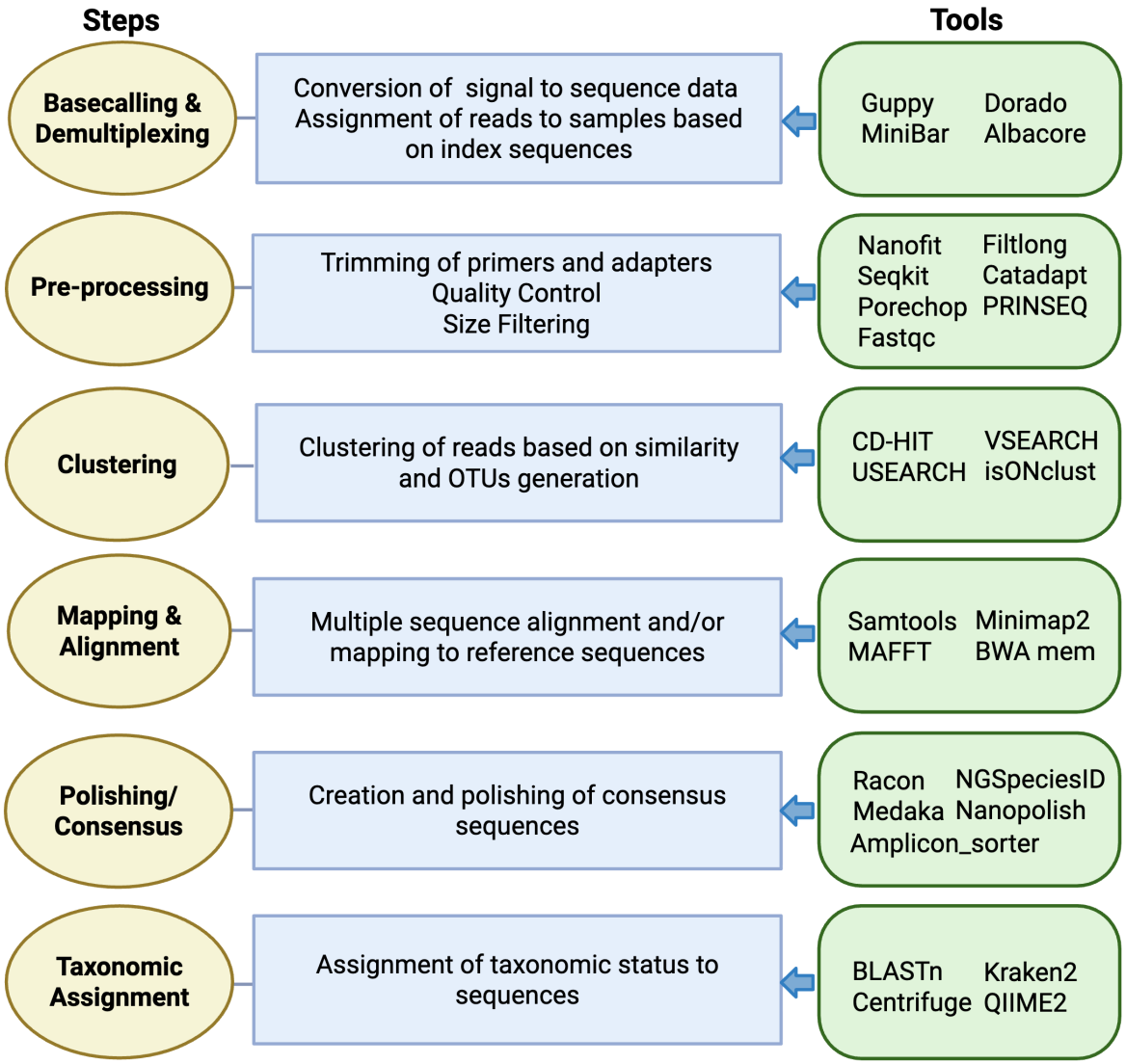
377 ASV analysis is performed by identifying sequence variations directly at the single-nucleotide level
378 without clustering, which retains more resolution (Overgaard et al., 2024). Extra processing steps like
379 dereplication and removal of chimeras can also be performed (Munian et al., 2024; Semmouri et al.,
380 2021; van der Reis et al., 2022). Taxonomic assignment of the identified OTUs or ASVs is performed
381 usually by BLAST (BLASTn, MegaBLAST) against a reference database containing barcode sequences
382 (Doorenspleet, Mailli, et al., 2025; van der Vorst et al., 2024). Overall, accurate taxonomic
383 assignment for metabarcoding datasets is linked to completeness of reference databases (McGee et
384 al., 2019; Zinger et al., 2019).

385 Few ONT metabarcoding bioinformatic pipelines are currently available (Supplementary Table 3).
386 Some of the most complete and commonly used are DECONA (Doorenspleet, Jansen, et al., 2025),
387 PIMENTA (van der Vorst et al., 2024) and MSI (Egeter et al., 2022), which perform similar overall
388 analysis, but with some variations. The DECONA pipeline is a simple command line tool which does
389 not require extensive bioinformatics background (Doorenspleet, Jansen, et al., 2025). DECONA starts
390 by read clustering, followed by alignment and assembly of clusters. After this, the consensus
391 sequences obtained are re-clustered, polished, and taxonomically assigned (Doorenspleet, Jansen, et
392 al., 2025). Similar steps are used for PIMENTA except re-clustering which is done with all samples
393 together, and without polishing of consensus sequences (van der Vorst et al., 2024), while MSI
394 performs polishing before clustering (Egeter et al., 2022). Other pipelines also exist, including: eNano,
395 designed as a "simple-to-understand" tool (Dierickx et al., 2024), ASHURE which is Python based and
396 performing multiple processing steps of metabarcoding (Baloğlu et al., 2021), and other approaches
397 for highly accurate consensus generation (NGSpeciesID) (Chang et al., 2024; Sahlin et al., 2021).

398 Method standardization for ONT data analysis is still lacking, which complicates the validation and
399 comparison of results between different studies (van der Vorst et al., 2024). Further development of
400 standardized bioinformatic pipelines, streamlining their use, and clear reporting of each step
401 undertaken will support ease of analysis and reproducibility of biodiversity studies in the future (Chang
402 et al., 2024; Harris et al., 2023; Kasmi et al., 2024; Truelove et al., 2019).

5c. PCR-Free (shotgun)

In shotgun sequencing, data originate from across the nuclear or organelle genome instead of a specific gene marker, and accurate assembly or species assignment can be a bioinformatic challenge. Furthermore, since this is a non-targeted approach, there is no specific size filtering, because the length of the random sequenced fragments is variable. To taxonomically assign reads in this case, mapping against reference genomes or organelles is performed (Bista et al., 2018; Linard et al., 2015; Quince et al., 2017). Most studies to date using PCR-free ONT sequencing target prokaryotes (Clark et al., 2025; Liu et al., 2024; Reska et al., 2024). Limited resources of eukaryotic reference genomes also make it difficult to apply this technique across many species, while higher sequencing effort than metabarcoding would be needed for standard detection. In Peel et al. (2019) a shotgun approach was used to analyse mixed pollen samples with a more cost effective and likely quantitative aspect than PCR-based. A pipeline (RevMet) was also developed which does not require assembly of a reference genome, but instead performs a genome skim, a low-coverage whole genome sequencing, which is used for identification (Peel et al., 2019).



417

418 **Figure 3**

419 **Bioinformatics tools.** Typical processing steps used for data analysis of long-read biodiversity datasets,
420 and suggested tools for each step. The order of processing steps may vary (sources listed in
421 Supplementary Table 3).

6. PacBio sequencing

PacBio sequencing was developed by Pacific Biosciences and is also known as “SMRT” sequencing due to performing Single-Molecule Real-Time sequencing (Rhoads & Au, 2015). The company became established in long-read sequencing in the early 2000s, marketing initially the PacBio RSI and II systems, and subsequently releasing the PacBio Sequel (I, II, and IIe) from 2015 onwards. Current instruments are the Revio, its bench top version the Vega, and various Sequel instruments still in use (Fig. 1B).

Sequencing is performed in SMRT cells, in Continuous Long Reads (CLR), and with the use of Circular Consensus Sequencing (CCS), which massively improves the final error rates. With CCS a consensus sequence is derived after multiple passes of each individual molecule, which makes it possible to convert noisy individual subreads into a highly accurate final read (Wenger et al., 2019) (Fig. 1B). With the introduction of High Fidelity (HiFi) reads and using CCS, the final accuracy exceeds 99% (Phred score 20), with a 10-30 Kb read length capability (Kucuk et al., 2023). The throughput of PacBio platforms increased largely with the introduction of the Revio which can yield up to 120 Gb per SMRT cell, compared to 30 Gb max throughput for the Sequel lle, while the Vega can yield approximately 60 Gb per SMRT cell.

7. Applications of PacBio in biodiversity studies

7.1 PacBio DNA barcoding

PacBio Sequel for DNA barcoding was used by Hebert et al. (2018) to test its multiplexing and read length capabilities, by sequencing >5,000 arthropod taxa in a single run. Furthermore, PacBio has been used for barcoding a wide diversity of fungi, including voucher specimens from herbaria and environmental lineages of parasitic fungi (Runnel et al., 2022; Wurzbacher et al., 2019). Higher success rate compared to Sanger was found especially in cases where polymorphic alleles were present, or when there was presence of contamination (Runnel et al., 2022).

In a comparison of DNA barcoding performance of Sanger, ONT, and PacBio platforms it was shown that both ONT and PacBio were suitable for DNA barcoding and biodiversity monitoring, while also suggesting that library preparation is less time consuming when using ONT (Cuber et al., 2023). Furthermore, high multiplexing and throughput can lower cost per barcode in comparison to Sanger sequencing (Hebert et al., 2018). The PacBio instruments are not currently widely used by individual labs due to their overall high annual service costs, as well as the cost of SMRT cells, which limits their presence mostly in core facilities (Hebert et al., 2018, 2025), though PacBio services are becoming more commonly available through sequencing providers.

7.2 PacBio metabarcoding of eDNA and community samples

Implementation of PacBio metabarcoding so far has been limited to few eukaryotic taxa, with most applications targeting fungi or microeukaryotes. Fungi represent a good case for long-read sequencing applications due to the known size variability and the complexity of markers used for analysis of fungal diversity such as ITS rRNA loci (Tedersoo et al., 2022). Initial fungal metabarcoding pipelines with PacBio RSII (CCS) (Heeger et al., 2018) or both RSII and Sequel (Kolaříková et al., 2021; Tedersoo et al.,

2018) aimed at generation of long amplicons (>2,500 bp) of the eukaryotic rRNA operon containing multiple commonly targeted regions (SSU, LSU, ITS). Sequencing of multiple long amplicons can aid in species identification, being particularly beneficial for use in phylogenetics, and to improve taxon recovery in cases of groups with low genetic (Tedersoo et al., 2018). Lichenized fungi in particular are interesting because individual specimens may harbour other fungi or mixed fungal communities making them difficult to sequence with Sanger sequencing. For these reasons Gueidan & Li (2022) and Gueidan et al. (2019) developed PacBio amplicon-based protocols for sequencing of lichen specimens from herbaria collections.

Furthermore, few studies to date have used PacBio to characterise microeukaryote communities from environmental or cultured samples (Jamy et al., 2020, 2022; Orr et al., 2018). For small unicellular organisms like protists, using PacBio can help bypass the need for cloning and Sanger sequencing which can be very laborious and costly (Orr et al., 2018). Furthermore, long amplicon data can be used to investigate evolutionary level questions, as was demonstrated by Jamy et al. (2022), where long-read amplicons aided in understanding habitat transitions across the eukaryote tree of life.

8. Sample prep and analysis of PacBio data

8.1 PacBio sample preparation

Extracting DNA for PacBio amplicon sequencing follows the same general considerations as with ONT. Methods vary based on application and as PacBio has been used mostly for fungal or microeukaryote sequencing the extraction types are more specific to these groups (Supplementary Table 1). In general, to take advantage of long-read capabilities, it is beneficial to consider using methods preserving long DNA fragments during extraction.

Sample preparation for PacBio-based DNA barcoding and metabarcoding involves initial PCR, with additional steps in order to attach barcodes for multiplexing (Fig. 2). For PacBio sequencing the double stranded DNA molecules are circularised after ligation with hairpin adaptors (Quail et al., 2012). This circularisation process enables repeated sequencing, generating multiple subreads as part of the CCS mode and HiFi sequencing (Cuber et al., 2023; Wenger et al., 2019).

For amplicon library preparation a SMRTbell library prep kit is used (current version 3.0) (Gueidan & Li, 2022). The amount of input DNA required for library preparation is related to the mean size of the target amplicons and varies for different instruments and chemistries. For example, for the latest Revio chemistry 50ng DNA are needed for amplicons <3 kb, increasing to >300 ng for amplicons >10 kb. For optimal results it is recommended to make calculations based on the specific approach and application of each study. Barcoding for multiplexing can be performed during the initial PCR following one or two-step protocols, with pooling prior to subsequent library prep steps. For the SMRTbell library preparation end-repair and A-tailing steps are also included and performed on the "barcoded" samples prior to adapter ligation. Finally, clean-up steps are performed at intervals and for the final library in accordance with the specific protocol (Fig. 2).

8.2 PacBio data analysis

PacBio amplicon data analysis includes steps and preprocessing similar to ONT (Fig. 3), though

individual pipelines have been developed to address various steps of the process that need further optimization. These include a denoising pipeline for allele phasing of PacBio data (Gueidan et al., 2019), or a phylogeny-aware taxonomic assignment method taking advantage of phylogenetic information available from long amplicons used to improve taxonomic classification (Jamy et al., 2020). Identification and removal of chimeric reads which are generated during PCR, also needs to be addressed with appropriate bioinformatic steps. A workflow developed by (Karst et al., 2021) uses unique molecular identifiers (UMIs) to reduce error rates and chimeras in long-read data (PacBio and ONT). UMIs, which are short sequences of 10-20 bases, are tagged to each template molecule enabling sorting and analysis of individual reads. This method was developed for long-read sequencing of microbial communities but could potentially be used for other amplicon applications (Overgaard et al., 2024). To support analyses of PacBio CCS amplicon data a pipeline was also developed based on the DADA2 algorithm and software, which was designed to determine ASVs with a high level resolution (single-nucleotide level) (Callahan et al., 2019).

9. Open data and reporting considerations

The principles of FAIR data reporting should be guiding all data driven scientific projects (Wilkinson et al., 2016), while when reporting biodiversity data these principles apply to sequencing data as well as tools, protocols, and workflows. During the development of the DNA barcoding and eDNA metabarcoding fields, data reporting has been evolving to reach community acceptable standards (Klymus et al., 2024). This reporting needs to be further extended to include details specific to long-read sequencing technologies. This is especially important due to the rapidly evolving nature of these technologies, which complicates even further method standardisation. It is suggested that the following information should be reported: For sequencing: the specific type of flow cell or SMRT cells and chemistry version, library preparation kit, adapter type (according to technology), and amount of input DNA and loaded library for sequencing. For data reporting: the read length profiles achieved, read coverage per sample, error rates, computational steps (and scripts) used for data analyses.

Most researchers submit raw data in public repositories such as INSDC which incorporates ENA and NCBI (Arita et al., 2021), nevertheless in that form they are difficult to process as multiple intermediate steps are required to get a usable form, such as OTUs or ASVs (Fig. 3). Submission of processed data from metabarcoding studies in public databases needs to be promoted by the community and is an important step for advancing reproducibility and data reuse (Berry et al., 2020). Current plans for integration of eDNA datasets in public repositories include databases such as GBIF (Frøslev et al., 2023). It is important to consider though that such datasets need to be filtered and monitored before release to ensure removal of inappropriate reads, like contaminants.

Open data platforms and projects can have a considerable impact on the advancement of open science. Reference databases such as the Barcode of Life Database (BOLD) database, are particularly important, acting as an open data repository and supporting storage, analysis, and coordination (BOLD.v4 (Ratnasingham et al., 2024). Large barcoding projects (e.g. BIOSCAN) also promote open data generation for DNA barcoding resources. The Darwin Tree of Life (DToL), which focuses on long-read whole genome sequencing, integrates DNA barcoding in their pipeline using it to support taxonomic identification of species for genome sequencing, and to populate barcode reference databases (Twyford et al., 2024).

Sequencing of whole genomes and organelles is necessary for future development, especially with the potential of long-read based biodiversity analysis. Such examples include taxon-specific organelle databases like MitoFish (fish mitogenomes) (Zhu et al., 2023), and ChloroplastDB (chloroplast genomes) (Cui et al., 2006). Additionally, initiatives such as the Earth BioGenome Project (EBP)(Lewin et al., 2018), DToL, and the European Reference Genomes Atlas (ERGA)(McCartney et al., 2024) are generating whole genome assemblies across the tree of life using long-read data. These initiatives which make their data immediately and openly available for use, constitute valuable resources for potential future genome-wide applications for biodiversity monitoring.

10. Limitations and future perspectives

The low concentration and degraded nature of eDNA (Goldberg et al., 2016; Thomsen et al., 2012), could hinder the use of long-read platforms, while also limiting certain types of applications such as shotgun sequencing. Environmental samples are known to contain inhibitors or contaminants which could damage ONT nanopores causing sequencing failure or reduced yield. PacBio platforms are less sensitive to such inhibitors due to different chemistry used though this has not been systematically explored for environmental samples. Careful optimization of DNA extraction and library prep protocols is needed to boost sequencing efficiency for long-read applications, along with eDNA grade quality control and clean-lab procedures (Alberdi et al., 2017; Bruce et al., 2021).

New developments in ONT chemistry and bioinformatics solutions have mostly dealt with high error rate issues. Particularly, improved basecalling algorithms, implementation of polishing steps, and improved workflows for clustering or classification have raised ONT data reliability to Illumina-like levels (Bludau et al., 2025; Chang et al., 2024; Sahlin et al., 2021). Similarly ONT and PacBio barcodes have been shown to present similar accuracy levels as Sanger and Illumina (Cuber et al., 2023; Srivathsan & Meier, 2024). Streamlining analysis and development of standardized bioinformatic pipelines for barcoding and metabarcoding will improve reliability of these applications and support future interstudy comparisons and reproducibility (van der Vorst et al., 2024).

Comparison of Illumina and ONT amplicon data demonstrated similar results for short amplicons (Egeter et al., 2022; Kasmi et al., 2024; Munian et al., 2024), but longer amplicons lacked detection efficiency (Doorenspleet, Jansen, et al., 2025; Maggini et al., 2024). This issue could be related to lower availability of long fragments in eDNA samples and a possible solution to improve long fragment detection could be to increase overall sequencing depth. Using longer fragments provides added value to the analysis by improving taxonomic assignment for closely related species or low intraspecific divergence (Doorenspleet, Jansen, et al., 2025; Tedersoo et al., 2018).

An overall comparison of PacBio vs ONT platforms reveals a number of interesting points. PacBio instruments require a higher financial investment for instrument purchase compared to the minimal setup required by ONT, which also offers faster turnaround times since individual labs can set-up and run ONT independently (Seah et al., 2020). Nevertheless, the new desktop Vega instrument from PacBio might be more accessible to individual labs, hence potentially also reducing turnaround times. A unique feature of the ONT MinION, its portability, allows sequencing in the field with a host of real-time sequencing applications which cannot currently be matched by other technologies.

Though ONT data quality is improving, PacBio read quality is already very high, especially since the introduction of HiFi reads (Chang et al., 2024; Kucuk et al., 2023). Throughput-wise the Revio presents a significant increase over previous instruments, reaching 120 Gb per SMRT cell, while a regular

MinION flow cell outputs a maximum of 48 Gb. Both ONT and PacBio offer high multiplexing capabilities (up to 10,000 samples) (Cuber et al., 2023) providing good value for money for large projects, especially for DNA barcoding and standardised metabarcoding.

ONT platforms offer more flexibility for small and medium sized projects with the Flongle and the MinION, while the GridION and PromethION instruments are more suitable for larger projects. Additionally, washing and reusing ONT flow cells could further reduce ONT costs. For PacBio, the Revio would be the instrument of choice for larger projects, while the Vega would be suitable for smaller projects. Though Vega has not been used in published biodiversity work yet, it could provide a potential solution for future applications. Selecting the most appropriate platform for a given project should take into consideration all the above factors as per the type of application, the nature and number of samples, and cost benefits for multiplexing.

Completeness of reference databases is an ongoing issue for DNA-based biodiversity monitoring applications (McGee et al., 2019; Zinger et al., 2019). Using short DNA barcode reference databases for taxonomic assignment of long-read data reduces the power of the application. In the future, reference databases need to also host longer fragments, and make use of outputs from whole genome and organelle sequencing efforts (Lewin et al., 2018; Theissinger et al., 2023; Zhu et al., 2023).

Sequencing whole mitogenomes from eDNA using long-read sequencers would represent a significant increase in data content provided from environmental samples. Despite eDNA's mostly degraded state, it has been shown that sequencing whole mitogenomes could be achieved through long-range amplification (Deiner, Renshaw, et al., 2017). This approach has been implemented mostly for community samples using shotgun sequencing, especially for arthropod communities (mito-metagenomics), nevertheless this experience could be further utilised for eDNA samples.

Expanding the use of shotgun sequencing could also resolve issues such as primer bias and abundance estimation biases (Alberdi et al., 2017; Bista et al., 2018), while offering exciting new opportunities for analysis. Shotgun sequencing could expand the scope of the data to wider areas of the nuclear and mitochondrial genome, moving beyond limited marker gene regions. This might make it possible to finally address population level questions. Currently population level work from eDNA has been limited mostly to species-specific assays (Parsons et al., 2018; Sigsgaard et al., 2016). In very few cases Illumina shotgun sequencing data have been used for population level analyses, such as for population inferences in cichlid fish from Lake Masoko (Liu et al., 2024), or for turtle population monitoring (Farrell et al., 2022). Shotgun sequencing effort would be higher and thus more expensive compared to metabarcoding if this approach was used for standard detection, nevertheless the additional gained information goes beyond detection. Additionally, utilization of whole genome data from consortia could contribute to further advancement of shotgun-based applications for biodiversity monitoring.

Finally, pairing long-read sequencing with other advancing technologies could be used to accelerate sample processing and data analysis. Employing robotic platforms for sample processing can boost operational speed and enable full use of the high multiplexing capabilities of long-read sequencing platforms. Such examples include development of high-throughput DNA extraction methods (Hebert et al., 2018), or metabarcoding sample processing workflows through liquid handling systems (Buchner et al., 2021). Furthermore, implementing Artificial Intelligence (AI) and cloud computing could dramatically improve analysis speed and accuracy for large datasets. Currently, AI has been used for ecosystem status predictions through eDNA by training machine learning algorithms (Keck et

al., 2023), and for improving annotation of eDNA reads (Flück et al., 2022), or DNA barcodes (Nanni et al., 2024). Overall, large potential exists for further development and use of AI solutions for DNA-based biodiversity monitoring coupled with long-read sequencing technologies.

11. Conclusions

The increasing impact of anthropogenic disturbances causing rapid loss of biodiversity presses the need for continued development of DNA-based applications for biodiversity monitoring. The recently developed field of long-read sequencing for biodiversity monitoring can offer improved capabilities and new solutions by introducing novel applications and complementing existing DNA-based monitoring tools. Both long-read sequencing platforms, ONT and PacBio, can offer solutions to different needs. Using long-read technologies could help push the limits of biodiversity detection to higher standards and address questions which were previously out of reach, such as population level assessments from environmental samples. Continuous introduction of new instruments and chemistries, along with improved workflows and bioinformatics tools keeps raising the bar of quantity and quality of the data offered. Nevertheless, standardisation of methodologies is still required, along with the united efforts from the research community in support of FAIR reporting, open data access and reproducibility, for future development.

Acknowledgements

This work was supported by funds awarded to IB by the Centre for Translational Biodiversity Genomics (LOEWE-TBG) through the program LOEWE–Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz of Hesse's Ministry of Higher Education, Research, and the Arts. We thank our colleagues for providing constructive feedback on the manuscript.

Author contributions

Iliana Bista: conceptualisation, leading design and writing. Iliana Bista, Alexandra Lino: wrote the manuscript.

Data availability

Not applicable

Conflict of interest

The authors have no conflict of interest to declare.

12. References

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2017). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147. https://doi.org/10.1111/2041-210X.12849

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968–3975. https://doi.org/10.1111/mec.14844

Arita, M., Karsch-Mizrachi, I., & Cochrane, G. (2021). The international nucleotide sequence database collaboration. *Nucleic Acids Research*, *49*(D1), D121–D124. https://doi.org/10.1093/nar/gkaa967

Baloğlu, B., Chen, Z., Elbrecht, V., Braukmann, T., MacDonald, S., & Steinke, D. (2021). A workflow for accurate metabarcoding using nanopore MinION sequencing. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13561

Berry, O., Jarman, S., Bissett, A., Hope, M., Paeper, C., Bessey, C., Schwartz, M. K., Hale, J., & Bunce, M. (2020). Making environmental DNA (eDNA) biodiversity records globally accessible. *Environmental DNA*. https://doi.org/10.1002/edn3.173

Bista, I., Carvalho, G. R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., Shokralla, S., Seymour, M., Bradley, D., Liu, S., Christmas, M., & Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*, *18*(5). https://doi.org/10.1111/1755-0998.12888

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Christmas, M., & Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, *8*, 14087. https://doi.org/10.1038/ncomms14087

Bludau, D., Sieber, G., Shah, M., Deep, A., Boenigk, J., & Beisser, D. (2025). Breaking the standard: can oxford nanopore technologies sequencing compete with Illumina in protistan amplicon studies? *Environmental DNA*, *7*(2). https://doi.org/10.1002/edn3.70084

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2022). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246. https://doi.org/10.1111/1755-0998.13512

Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S. B., Krstic, P. S., Lindsay, S., Ling, X. S., Mastrangelo, C. H., Meller, A., Oliver, J. S., Pershin, Y. V., Ramsey, J. M., … Schloss, J. A. (2008). The potential and challenges of nanopore sequencing. *Nature Biotechnology*, *26*(10), 1146–1153. https://doi.org/10.1038/nbt.1495

Bruce, K., Blackman, R., Bourlat, S. J., Hellstrom, A. M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V., Hänfling, B., Leese, F., Mächler, E., Mahon, A. R., Meissner, K., Panksep, K., … Deiner, K. (2021). *A practical guide to DNA-based methods for biodiversity assessment*. Pensoft Publishers. https://doi.org/10.3897/ab.e68634

Buchner, D., Macher, T.-H., Beermann, A. J., Werner, M.-T., & Leese, F. (2021). Standardized high-throughput biomonitoring using DNA metabarcoding: Strategies for the adoption of automated liquid handlers. *Environmental Science and Ecotechnology*, *8*, 100122. https://doi.org/10.1016/j.ese.2021.100122

Callahan, B. J., Wong, J., Heiner, C., Oh, S., Theriot, C. M., Gulati, A. S., McGill, S. K., & Dougherty, M. K. (2019). High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Research*, *47*(18), e103. https://doi.org/10.1093/nar/gkz569

Chang, J. J. M., Ip, Y. C. A., Neo, W. L., Mowe, M. A. D., Jaafar, Z., & Huang, D. (2024). Primed and ready: nanopore metabarcoding can now recover highly accurate consensus barcodes that are generally indel-free. *BMC Genomics*, *25*(1), 842. https://doi.org/10.1186/s12864-024-10767-4

Clark, M. S., Gregson, B. H., Greco, C., Nair, H. P., Clark, M., Evans, C., Hughes, K. A., Hwengwere, K., Leung, M., Peck, L. S., Walker, C. A., & Chow, W. (2025). Assessing the impact of sewage and wastewater on antimicrobial resistance in nearshore Antarctic biofilms and sediments. *Environmental Microbiome*, *20*(1), 9. https://doi.org/10.1186/s40793-025-00671-z

Cuber, P., Chooneea, D., Geeves, C., Salatino, S., Creedy, T. J., Griffin, C., Sivess, L., Barnes, I., Price, B., & Misra, R. (2023). Comparing the accuracy and efficiency of third generation sequencing technologies, Oxford Nanopore Technologies, and Pacific Biosciences, for DNA barcode sequencing applications. *Ecological Genetics and Genomics*, *28*, 100181. https://doi.org/10.1016/j.egg.2023.100181

Cui, L., Veeraraghavan, N., Richter, A., Wall, K., Jansen, R. K., Leebens-Mack, J., Makalowska, I., & dePamphilis, C. W. (2006). Chloroplastdb: the chloroplast genome database. *Nucleic Acids Research*, *34*(Database issue), D692-6. https://doi.org/10.1093/nar/gkj055

Deamer, D., Akeson, M., & Branton, D. (2016). Three decades of nanopore sequencing. *Nature Biotechnology*, *34*(5), 518–524. https://doi.org/10.1038/nbt.3423

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., & Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.12836

Dierickx, G., Tondeleir, L., Asselman, P., Vandekerkhove, K., & Verbeken, A. (2024). What quality suffices for Nanopore metabarcoding? Reconsidering methodology and ectomycorrhizae in decaying Fagus sylvatica bark as case study. *Journal of Fungi (Basel, Switzerland)*, *10*(10). https://doi.org/10.3390/jof10100708

Doorenspleet, K., Jansen, L., Oosterbroek, S., Kamermans, P., Bos, O., Wurz, E., Murk, A., & Nijland, R. (2025). The long and the short of it: Nanopore-based eDNA metabarcoding of marine vertebrates works; Sensitivity and species-level assignment depend on amplicon lengths. *Molecular Ecology Resources*, e14079. https://doi.org/10.1111/1755-0998.14079

Doorenspleet, K., Mailli, A. A., van der Hoorn, B. B., Beentjes, K. K., De Backer, A., Derycke, S., Murk, A. J., Reiss, H., & Nijland, R. (2025). Advancing molecular macrobenthos biodiversity monitoring: a comparison between Oxford Nanopore and Illumina based metabarcoding and metagenomics. *PeerJ*, *13*, e19158. https://doi.org/10.7717/peerj.19158

Egeter, B., Veríssimo, J., Lopes-Lima, M., Chaves, C., Pinto, J., Riccardi, N., Beja, P., & Fonseca, N. A. (2022). Speeding up the detection of invasive bivalve species using environmental DNA: A Nanopore and Illumina sequencing comparison. *Molecular Ecology Resources*, *22*(6), 2232–2247. https://doi.org/10.1111/1755-0998.13610

Farrell, J. A., Whitmore, L., Mashkour, N., Rollinson Ramia, D. R., Thomas, R. S., Eastman, C. B., Burkhalter, B., Yetsko, K., Mott, C., Wood, L., Zirkelbach, B., Meers, L., Kleinsasser, P., Stock, S., Libert, E., Herren, R., Eastman, S., Crowder, W., Bovery, C., … Duffy, D. J. (2022). Detection and population genomics of sea turtle species via noninvasive environmental DNA analysis of nesting beach sand tracks and oceanic water. *Molecular Ecology Resources*, *22*(7), 2471–2493. https://doi.org/10.1111/1755-0998.13617

Flück, B., Mathon, L., Manel, S., Valentini, A., Dejean, T., Albouy, C., Mouillot, D., Thuiller, W., Murienne, J., Brosse, S., & Pellissier, L. (2022). Applying convolutional neural networks to speed up environmental DNA annotation in a highly diverse ecosystem. *Scientific Reports*, *12*(1), 10247. https://doi.org/10.1038/s41598-022-13412-w

Frøslev, T., Jeppesen, T., Robertson, T., & Schigel, D. (2023). eDNA metabarcoding data in the Global Biodiversity Information Facility (GBIF). *Biodiversity Information Science and Standards*, *7*. https://doi.org/10.3897/biss.7.110672

Gajski, D., Wolff, J. O., Melcher, A., Weber, S., Prost, S., Krehenwinkel, H., & Kennedy, S. R. (2024). Facilitating taxonomy and phylogenetics: An informative and cost-effective protocol integrating long amplicon PCRs and third-generation sequencing. *Molecular Phylogenetics and Evolution*, *192*, 107988. https://doi.org/10.1016/j.ympev.2023.107988

Gaonkar, C. C., & Campbell, L. (2024). A full-length 18S ribosomal DNA metabarcoding approach for determining protist community diversity using Nanopore sequencing. *Ecology and Evolution*, *14*(4), e11232. https://doi.org/10.1002/ece3.11232

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., & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.12595

Gueidan, C., Elix, J. A., McCarthy, P. M., Roux, C., Mallen-Cooper, M., & Kantvilas, G. (2019). PacBio amplicon sequencing for metabarcoding of mixed DNA samples from lichen herbarium specimens. *MycoKeys*, *53*, 73–91. https://doi.org/10.3897/mycokeys.53.34761

Gueidan, C., & Li, L. (2022). A long-read amplicon approach to scaling up the metabarcoding of lichen herbarium specimens. *MycoKeys*, *86*, 195–212. https://doi.org/10.3897/mycokeys.86.77431

Gygax, D., Ramirez, S., Chibesa, M., Simpamba, T., Riffel, M., Riffel, T., Zulu, G., Srivathsan, A., Nijland, R., & Urban, L. (2025). Evaluation of nanopore sequencing for increasing accessibility of eDNA studies in biodiverse countries. *BioRxiv*. https://doi.org/10.1101/2025.04.21.649756

Hakimzadeh, A., Abdala Asbun, A., Albanese, D., Bernard, M., Buchner, D., Callahan, B., Caporaso, J. G., Curd, E., Djemiel, C., Brandström Durling, M., Elbrecht, V., Gold, Z., Gweon, H. S., Hajibabaei, M., Hildebrand, F., Mikryukov, V., Normandeau, E., Özkurt, E., M Palmer, J., … Anslan, S. (2024). A pile of pipelines: An overview of the bioinformatics software for metabarcoding data analyses. *Molecular Ecology Resources*, *24*(5), e13847. https://doi.org/10.1111/1755-0998.13847

Harris, S. E., Whitehurst, A., Buehrer, M., Lonker, S., Veverka, B., & Nagy, C. (2023). Efficient multiplexing of pollinator metabarcodes using Oxford Nanopore MinION sequencing: insights for meadow management from floral environmental DNA. *BioRxiv*. https://doi.org/10.1101/2023.12.15.571580

Hartop, E., Lee, L., Srivathsan, A., Jones, M., Peña-Aguilera, P., Ovaskainen, O., Roslin, T., & Meier, R. (2024). Resolving biology's dark matter: species richness, spatiotemporal distribution, and community composition of a dark taxon. *BMC Biology*, *22*(1), 215. https://doi.org/10.1186/s12915-024-02010-z

Hebert, Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., Janzen, D. H., Hallwachs, W., Naik, S., Sones, J. E., & Zakharov, E. V. (2018). A Sequel to Sanger: amplicon sequencing that scales. *BMC Genomics*, *19*(1), 219. https://doi.org/10.1186/s12864-018-4611-3

Hebert, Floyd, R., Jafarpour, S., & Prosser, S. W. J. (2025). Barcode 100K specimens: in a single Nanopore run. *Molecular Ecology Resources*, *25*(1), e14028. https://doi.org/10.1111/1755-0998.14028

Heeger, F., Bourne, E. C., Baschien, C., Yurkov, A., Bunk, B., Spröer, C., Overmann, J., Mazzoni, C. J., & Monaghan, M. T. (2018). Long-read DNA metabarcoding of ribosomal RNA in the analysis of fungi

from aquatic environments. *Molecular Ecology Resources*, *18*(6), 1500–1514. https://doi.org/10.1111/1755-0998.12937

Ho, J. K. I., Puniamoorthy, J., Srivathsan, A., & Meier, R. (2020). MinION sequencing of seafood in Singapore reveals creatively labelled flatfishes, confused roe, pig DNA in squid balls, and phantom crustaceans. *Food Control*, *112*, 107144. https://doi.org/10.1016/j.foodcont.2020.107144

Huggins, L. G., Colella, V., Young, N. D., & Traub, R. J. (2024). Metabarcoding using nanopore long-read sequencing for the unbiased characterization of apicomplexan haemoparasites. *Molecular Ecology Resources*, *24*(2), e13878. https://doi.org/10.1111/1755-0998.13878

Ip, Y. C. A., Chang, J. J. M., Oh, R. M., Quek, Z. B. R., Chan, Y. K. S., Bauman, A. G., & Huang, D. (2023). Seq' and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs. *Molecular Ecology*, *32*(23), 6223–6242. https://doi.org/10.1111/mec.16568

Jamy, M., Biwer, C., Vaulot, D., Obiol, A., Jing, H., Peura, S., Massana, R., & Burki, F. (2022). Global patterns and rates of habitat transitions across the eukaryotic tree of life. *Nature Ecology & Evolution*, *6*(10), 1458–1470. https://doi.org/10.1038/s41559-022-01838-4

Jamy, M., Foster, R., Barbera, P., Czech, L., Kozlov, A., Stamatakis, A., Bending, G., Hilton, S., Bass, D., & Burki, F. (2020). Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. *Molecular Ecology Resources*, *20*(2), 429–443. https://doi.org/10.1111/1755-0998.13117

Karst, S. M., Ziels, R. M., Kirkegaard, R. H., Sørensen, E. A., McDonald, D., Zhu, Q., Knight, R., & Albertsen, M. (2021). High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nature Methods*, *18*(2), 165–169. https://doi.org/10.1038/s41592-020-01041-y

Kasmi, Y., Neumann, H., Haslob, H., Blancke, T., Möckel, B., Postel, U., & Hanel, R. (2024). Comparative analysis of bottom trawl and nanopore sequencing in fish biodiversity assessment: The sylt outer reef example. *Marine Environmental Research*, *199*, 106602. https://doi.org/10.1016/j.marenvres.2024.106602

Keck, F., Brantschen, J., & Altermatt, F. (2023). A combination of machine-learning and eDNA reveals the genetic signature of environmental change at the landscape levels. *Molecular Ecology*, *32*(17), 4791–4800. https://doi.org/10.1111/mec.17073

Kelly, R. P., Port, J. A., Yamahara, K. M., Martone, R. G., Lowell, N., Thomsen, P. F., Mach, M. E., Bennett, M., Prahler, E., Caldwell, M. R., & Crowder, L. B. (2014). Environmental monitoring. Harnessing DNA to improve environmental management. *Science*, *344*(6191), 1455–1456. https://doi.org/10.1126/science.1251156

Kim, C., Pongpanich, M., & Porntaveetus, T. (2024). Unraveling metagenomics through long-read sequencing: a comprehensive review. *Journal of Translational Medicine*, *22*(1), 111. https://doi.org/10.1186/s12967-024-04917-1

Klymus, K. E., Baker, J. D., Abbott, C. L., Brown, R. J., Craine, J. M., Gold, Z., Hunter, M. E., Johnson, M. D., Jones, D. N., Jungbluth, M. J., Jungbluth, S. P., Lor, Y., Maloy, A. P., Merkes, C. M., Noble, R., Patin, N. V., Sepulveda, A. J., Spear, S. F., Steele, J. A., … Theroux, S. (2024). The MIEM guidelines: Minimum information for reporting of environmental metabarcoding data. *Metabarcoding and Metagenomics*, *8*. https://doi.org/10.3897/mbmg.8.128689

Knot, I. E., Zouganelis, G. D., Weedall, G. D., Wich, S. A., & Rae, R. (2020). DNA barcoding of nematodes using the minion. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.00100

Koblmüller, S., Resl, P., Klar, N., Bauer, H., Zangl, L., & Hahn, C. (2024). DNA Barcoding for species identification of moss-dwelling invertebrates: Performance of Nanopore sequencing and coverage in reference database. *Diversity*, *16*(4), 196. https://doi.org/10.3390/d16040196

Koda, S. A., McCauley, M., Farrell, J. A., Duffy, I. J., Duffy, F. G., Loesgen, S., Whilde, J., & Duffy, D. J. (2023). A novel eDNA approach for rare species monitoring: Application of long-read shotgun sequencing to *Lynx rufus* soil pawprints. *Biological Conservation*, *287*, 110315. https://doi.org/10.1016/j.biocon.2023.110315

Kolaříková, Z., Slavíková, R., Krüger, C., Krüger, M., & Kohout, P. (2021). PacBio sequencing of Glomeromycota rDNA: a novel amplicon covering all widely used ribosomal barcoding regions and its applicability in taxonomy and ecology of arbuscular mycorrhizal fungi. *The New Phytologist*, *231*(1), 490–499. https://doi.org/10.1111/nph.17372

Krehenwinkel, H., Pomerantz, A., Henderson, J. B., Kennedy, S. R., Lim, J. Y., Swamy, V., Shoobridge, J. D., Graham, N., Patel, N. H., Gillespie, R. G., & Prost, S. (2019). Nanopore sequencing of long ribosomal DNA amplicons enables portable and simple biodiversity assessments with high phylogenetic resolution across broad taxonomic scale. *GigaScience*, *8*(5). https://doi.org/10.1093/gigascience/giz006

Krehenwinkel, H., Pomerantz, A., & Prost, S. (2019). Genetic biomonitoring and biodiversity assessment using portable sequencing technologies: current uses and future directions. *Genes*, *10*(11). https://doi.org/10.3390/genes10110858

Kucuk, E., van der Sanden, B. P. G. H., O'Gorman, L., Kwint, M., Derks, R., Wenger, A. M., Lambert, C., Chakraborty, S., Baybayan, P., Rowell, W. J., Brunner, H. G., Vissers, L. E. L. M., Hoischen, A., & Gilissen, C. (2023). Comprehensive de novo mutation discovery with HiFi long-read sequencing. *Genome Medicine*, *15*(1), 34. https://doi.org/10.1186/s13073-023-01183-6

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., Durbin, R., Edwards, S. V., Forest, F., Gilbert, M. T. P., Goldstein, M. M., Grigoriev, I. V., Hackett, K. J.,

Haussler, D., Jarvis, E. D., Johnson, W. E., Patrinos, A., Richards, S., Castilla-Rubio, J. C., … Zhang, G. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(17), 4325–4333. https://doi.org/10.1073/pnas.1720115115

Linard, B., Crampton-Platt, A., Gillett, C. P. D. T., Timmermans, M. J. T. N., & Vogler, A. P. (2015). Metagenome skimming of insect specimen pools: potential for comparative genomics. *Genome Biology and Evolution*, *7*(6), 1474–1489. https://doi.org/10.1093/gbe/evv086

Liu, Z., Kishe, M. A., Gabagambi, N. P., Shechonge, A. H., Ngatunga, B. P., Smith, K., Saxon, A. D., Hudson, A. G., Linderoth, T., Turner, G. F., Collins, R. A., & Genner, M. J. (2024). Nuclear environmental DNA resolves fine-scale population genetic structure in an aquatic habitat. *IScience*, *27*(1), 108669. https://doi.org/10.1016/j.isci.2023.108669

Loh, R. K., Tan, T. R. H., Yeo, H., Yeoh, T. X., Lee, T. T. M., Kutty, S. N., & Puniamoorthy, N. (2024). Metabarcoding mosquitoes: MinION sequencing of bulk samples gives accurate species profiles for vector surveillance (Culicidae). *Frontiers in Tropical Diseases*, *5*. https://doi.org/10.3389/fitd.2024.1223435

MacKenzie, M., & Argyropoulos, C. (2023). An introduction to nanopore sequencing: past, present, and future considerations. *Micromachines*, *14*(2). https://doi.org/10.3390/mi14020459

Maestri, S., Cosentino, E., Paterno, M., Freitag, H., Garces, J. M., Marcolungo, L., Alfano, M., Njunjić, I., Schilthuizen, M., Slik, F., Menegon, M., Rossato, M., & Delledonne, M. (2019). A rapid and accurate MinION-based workflow for tracking species biodiversity in the field. *Genes*, *10*(6). https://doi.org/10.3390/genes10060468

Maggini, S., Jacobsen, M. W., Urban, P., Hansen, B. K., Kielgast, J., Bekkevold, D., Jardim, E., Martinsohn, J. T., Carvalho, G. R., Nielsen, E. E., & Papadopulos, A. S. T. (2024). Nanopore environmental DNA sequencing of catch water for estimating species composition in demersal bottom trawl fisheries. *Environmental DNA*, *6*(3). https://doi.org/10.1002/edn3.555

McCartney, A. M., Formenti, G., Mouton, A., De Panis, D., Marins, L. S., Leitão, H. G., Diedericks, G., Kirangwa, J., Morselli, M., Salces-Ortiz, J., Escudero, N., Iannucci, A., Natali, C., Svardal, H., Fernández, R., De Pooter, T., Joris, G., Strazisar, M., Wood, J. M. D., … et al. (2024). The European Reference Genome Atlas: piloting a decentralised approach to equitable biodiversity genomics. *Npj Biodiversity*, *3*(1), 28. https://doi.org/10.1038/s44185-024-00054-6

McGee, K. M., Robinson, C. V., & Hajibabaei, M. (2019). Gaps in DNA-based biomonitoring across the globe. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00337

Menegon, M., Cantaloni, C., Rodriguez-Prieto, A., Centomo, C., Abdelfattah, A., Rossato, M., Bernardi, M., Xumerle, L., Loader, S., & Delledonne, M. (2017). On site DNA barcoding by nanopore sequencing. *Plos One*, *12*(10), e0184741. https://doi.org/10.1371/journal.pone.0184741

902 Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S.,
903 Yamanaka, H., Araki, H., Kondoh, M., & Iwasaki, W. (2015). MiFish, a set of universal PCR primers
904 for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical
905 marine species. *Royal Society Open Science*, *2*(7), 150088. https://doi.org/10.1098/rsos.150088

906 Munian, K., Ramli, F. F., Othman, N., Mahyudin, N. A. A., Sariyati, N. H., Abdullah-Fauzi, N. A. F.,
907 Haris, H., Ilham-Norhakim, M. L., & Abdul-Latiff, M. A. B. (2024). Environmental DNA
908 metabarcoding of freshwater fish in Malaysian tropical rivers using short-read nanopore
909 sequencing as a potential biomonitoring tool. *Molecular Ecology Resources*, *24*(4), e13936.
910 https://doi.org/10.1111/1755-0998.13936

911 Nanni, L., Cuza, D., & Brahnam, S. (2024). AI-Powered Biodiversity Assessment: Species Classification
912 via DNA Barcoding and Deep Learning. *Technologies*, *12*(12), 240.
913 https://doi.org/10.3390/technologies12120240

914 Nousias, O., Duffy, F. G., Duffy, I. J., Whilde, J., & Duffy, D. J. (2024). Long-read Nanopore shotgun
915 eDNA sequencing for river biodiversity, pollution and environmental health monitoring. *BioRxiv*.
916 https://doi.org/10.1101/2024.11.01.621618

917 Nygaard, A. B., Tunsjø, H. S., Meisal, R., & Charnock, C. (2020). A preliminary study on the potential
918 of Nanopore MinION and Illumina MiSeq 16S rRNA gene sequencing to characterize building-dust
919 microbiomes. *Scientific Reports*, *10*(1), 3209. https://doi.org/10.1038/s41598-020-59771-0

920 Orr, R. J. S., Zhao, S., Klaveness, D., Yabuki, A., Ikeda, K., Watanabe, M. M., & Shalchian-Tabrizi, K.
921 (2018). Enigmatic Diphyllatea eukaryotes: culturing and targeted PacBio RS amplicon sequencing
922 reveals a higher order taxonomic diversity and global distribution. *BMC Evolutionary Biology*,
923 *18*(1), 115. https://doi.org/10.1186/s12862-018-1224-z

924 Overgaard, C. K., Jamy, M., Radutoiu, S., Burki, F., & Dueholm, M. K. D. (2024). Benchmarking long-
925 read sequencing strategies for obtaining ASV-resolved rRNA operons from environmental
926 microeukaryotes. *Molecular Ecology Resources*, *24*(7), e13991. https://doi.org/10.1111/1755-
927 0998.13991

928 Parsons, K. M., Everett, M., Dahlheim, M., & Park, L. (2018). Water, water everywhere:
929 environmental DNA can unlock population structure in elusive marine species. *Royal Society Open*
930 *Science*, *5*(8), 180537. https://doi.org/10.1098/rsos.180537

931 Peel, N., Dicks, L. V., Clark, M. D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R. G., Leggett, R.
932 M., & Yu, D. W. (2019). Semi-quantitative characterisation of mixed pollen samples using MinION
933 sequencing and Reverse Metagenomics (RevMet). *Methods in Ecology and Evolution*.
934 https://doi.org/10.1111/2041-210X.13265

935 Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amorós,
936 C. L., Salazar-Valenzuela, D., & Prost, S. (2018). Real-time DNA barcoding in a rainforest using

Nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, *7*(4). https://doi.org/10.1093/gigascience/giy033

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P., & Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, *13*, 341. https://doi.org/10.1186/1471-2164-13-341

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*(9), 833–844. https://doi.org/10.1038/nbt.3935

Rang, F. J., Kloosterman, W. P., & de Ridder, J. (2018). From squiggle to basepair: computational approaches for improving Nanopore sequencing read accuracy. *Genome Biology*, *19*(1), 90. https://doi.org/10.1186/s13059-018-1462-9

Ratnasingham, S., Wei, C., Chan, D., Agda, J., Agda, J., Ballesteros-Mejia, L., Boutou, H. A., El Bastami, Z. M., Ma, E., Manjunath, R., Rea, D., Ho, C., Telfer, A., McKeowan, J., Rahulan, M., Steinke, C., Dorsheimer, J., Milton, M., & Hebert, P. D. N. (2024). BOLD v4: A Centralized bioinformatics platform for DNA-Based biodiversity data. *Methods in Molecular Biology*, *2744*, 403–441. https://doi.org/10.1007/978-1-0716-3581-0_26

Reska, T., Pozdniakova, S., Borràs, S., Perlas, A., Sauerborn, E., Cañas, L., Schloter, M., Rodó, X., Wang, Y., Winkler, B., Schnitzler, J.-P., & Urban, L. (2024). Air monitoring by Nanopore sequencing. *ISME Communications*, *4*(1), ycae099. https://doi.org/10.1093/ismeco/ycae099

Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics / Beijing Genomics Institute*, *13*(5), 278–289. https://doi.org/10.1016/j.gpb.2015.08.002

Runnel, K., Abarenkov, K., Copoț, O., Mikryukov, V., Kõljalg, U., Saar, I., & Tedersoo, L. (2022). DNA barcoding of fungal specimens using PacBio long-read high-throughput sequencing. *Molecular Ecology Resources*, *22*(8), 2871–2879. https://doi.org/10.1111/1755-0998.13663

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, *17*, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sahlin, K., Lim, M. C. W., & Prost, S. (2021). NGSpeciesID: DNA barcode and amplicon consensus generation from long-read sequencing data. *Ecology and Evolution*, *11*(3), 1392–1398. https://doi.org/10.1002/ece3.7146

Seah, A., Lim, M. C. W., McAloose, D., Prost, S., & Seimon, T. A. (2020). MinION-based DNA Barcoding of preserved and non-invasively collected wildlife samples. *Genes*, *11*(4). https://doi.org/10.3390/genes11040445

973 Semmouri, I., De Schamphelaere, K. A. C., Willemse, S., Vandegehuchte, M. B., Janssen, C. R., &
974 Asselman, J. (2021). Metabarcoding reveals hidden species and improves identification of marine
975 zooplankton communities in the North Sea. *ICES Journal of Marine Science*, *78*(9), 3411–3427.
976 https://doi.org/10.1093/icesjms/fsaa256

977 Shum, P., Cusa, M., Prasetyo, A., & Mariani, S. (2024). Nanopore sequencing facilitates screening of
978 diversity and provenance of seafood and marine wildlife. *Food Control*, *161*, 110382.
979 https://doi.org/10.1016/j.foodcont.2024.110382

980 Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., Pedersen,
981 M. W., Jaidah, M. A., Orlando, L., Willerslev, E., Møller, P. R., & Thomsen, P. F. (2016). Population
982 characteristics of a large whale shark aggregation inferred from seawater environmental DNA.
983 *Nature Ecology & Evolution*, *1*(1), 4. https://doi.org/10.1038/s41559-016-0004

984 Simon, S. A., Schmidt, K., Griesdorn, L., Soares, A. R., Bornemann, T. L. V., & Probst, A. J. (2023).
985 Dancing the Nanopore limbo - Nanopore metagenomics from small DNA quantities for bacterial
986 genome reconstruction. *BMC Genomics*, *24*(1), 727. https://doi.org/10.1186/s12864-023-09853-w

987 Srivathsan, A., Baloğlu, B., Wang, W., Tan, W. X., Bertrand, D., Ng, A. H. Q., Boey, E. J. H., Koh, J. J. Y.,
988 Nagarajan, N., & Meier, R. (2018). A MinION™-based pipeline for fast and cost-effective DNA
989 barcoding. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12890

990 Srivathsan, A., Feng, V., Suárez, D., Emerson, B., & Meier, R. (2024). ONTbarcoder 2.0: rapid species
991 discovery and identification with real-time barcoding facilitated by Oxford Nanopore R10.4.
992 *Cladistics*, *40*(2), 192–203. https://doi.org/10.1111/cla.12566

993 Srivathsan, A., Hartop, E., Puniamoorthy, J., Lee, W. T., Kutty, S. N., Kurina, O., & Meier, R. (2019).
994 Rapid, large-scale species discovery in hyperdiverse taxa using 1D MinION sequencing. *BMC*
995 *Biology*, *17*(1), 96. https://doi.org/10.1186/s12915-019-0706-9

996 Srivathsan, A., Lee, L., Katoh, K., Hartop, E., Kutty, S. N., Wong, J., Yeo, D., & Meier, R. (2021).
997 ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for
998 everyone. *BMC Biology*, *19*(1), 217. https://doi.org/10.1186/s12915-021-01141-x

999 Srivathsan, A., & Meier, R. (2024). Scalable, Cost-Effective, and Decentralized DNA Barcoding with
1000 Oxford Nanopore Sequencing. *Methods in Molecular Biology*, *2744*, 223–238.
1001 https://doi.org/10.1007/978-1-0716-3581-0_14

1002 Stevens, B. M., Creed, T. B., Reardon, C. L., & Manter, D. K. (2023). Comparison of Oxford Nanopore
1003 Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil.
1004 *Scientific Reports*, *13*(1), 9323. https://doi.org/10.1038/s41598-023-36101-8

1005 Stoeck, T., Katzenmeier, S. N., Breiner, H.-W., & Rubel, V. (2024). Nanopore duplex sequencing as an
1006 alternative to Illumina MiSeq sequencing for eDNA-based biomonitoring of coastal aquaculture
1007 impacts. *Metabarcoding and Metagenomics*, *8*. https://doi.org/10.3897/mbmg.8.121817

1008 Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-
1009 generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–
1010 2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

1011 Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R. H., Kennedy, P. G., Yang, T., Anslan, S., & Mikryukov,
1012 V. (2022). Best practices in metabarcoding of fungi: From experimental design to results.
1013 *Molecular Ecology*, *31*(10), 2769–2795. https://doi.org/10.1111/mec.16460

1014 Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). PacBio metabarcoding of Fungi and other
1015 eukaryotes: errors, biases and perspectives. *The New Phytologist*, *217*(3), 1370–1385.
1016 https://doi.org/10.1111/nph.14776

1017 Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., Bombarely, A., Crottini,
1018 A., Gallo, G. R., Godoy, J. A., Jentoft, S., Malukiewicz, J., Mouton, A., Oomen, R. A., Paez, S.,
1019 Palsbøll, P. J., Pampoulie, C., Ruiz-López, M. J., Secomandi, S., … European Reference Genome
1020 Atlas Consortium. (2023). How genomics can help biodiversity conservation. *Trends in Genetics*,
1021 *39*(7), 545–559. https://doi.org/10.1016/j.tig.2023.01.005

1022 Theologidis, I., Karamitros, T., Vichou, A.-E., & Kizis, D. (2023). Nanopore sequencing metabarcoding
1023 for identification of phytopathogenic and endophytic fungi in olive (*Olea europaea*) twigs. *Journal*
1024 *of Fungi (Basel, Switzerland)*, *9*(11). https://doi.org/10.3390/jof9111119

1025 Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L., &
1026 Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA.
1027 *Molecular Ecology*, *21*(11), 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

1028 Toxqui Rodríguez, M. del S., Vanhollebeke, J., & Derycke, S. (2023). Evaluation of DNA
1029 metabarcoding using Oxford Nanopore sequencing for authentication of mixed seafood products.
1030 *Food Control*, *145*, 109388. https://doi.org/10.1016/j.foodcont.2022.109388

1031 Truelove, N. K., Andruszkiewicz, E. A., & Block, B. A. (2019). A rapide DNA method for detecting
1032 white sharks in the open ocean. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-
1033 210X.13201

1034 Twyford, A. D., Beasley, J., Barnes, I., Allen, H., Azzopardi, F., Bell, D., Blaxter, M. L., Broad, G.,
1035 Campos-Dominguez, L., Choonea, D., Crowley, L., Cuber, P., Cunliffe, M., Dombrowski, A., Douglas,
1036 B., Forrest, L. L., Gaya, E., Greeves, C., Griffin, C., … Darwin Tree of Life Consortium. (2024). A DNA
1037 barcoding framework for taxonomic verification in the Darwin Tree of Life Project. *Wellcome Open*
1038 *Research*, *9*, 339. https://doi.org/10.12688/wellcomeopenres.21143.1

1039 Urban, L., Holzer, A., Baronas, J. J., Hall, M. B., Braeuninger-Weimer, P., Scherm, M. J., Kunz, D. J.,
1040 Perera, S. N., Martin-Herranz, D. E., Tipper, E. T., Salter, S. J., & Stammnitz, M. R. (2021).
1041 Freshwater monitoring by nanopore sequencing. *ELife*, *10*. https://doi.org/10.7554/eLife.61504

Urban, L., Miller, A. K., Eason, D., Vercoe, D., Shaffer, M., Wilkinson, S. P., Jeunen, G.-J., Gemmell, N. J., & Digby, A. (2023). Non-invasive real-time genomic monitoring of the critically endangered kākāpō. *ELife*, *12*. https://doi.org/10.7554/eLife.84553

van der Reis, A. L., Beckley, L. E., Olivar, M. P., & Jeffs, A. G. (2022). Nanopore short-read sequencing: A quick, cost-effective and accurate method for DNA metabarcoding. *Environmental DNA*. https://doi.org/10.1002/edn3.374

van der Vorst, V., Thijssen, M., Fronen, B. J., de Groot, A., Maathuis, M. A. M., Nijhuis, E., Polling, M., Stassen, J., Voorhuijzen-Harink, M. M., & Jak, R. (2024). PIMENTA: Pipeline for metabarcoding through Nanopore technology used for authentication. *BioRxiv*. https://doi.org/10.1101/2024.02.14.580249

Vasilita, C., Feng, V., Hansen, A. K., Hartop, E., Srivathsan, A., Struijk, R., & Meier, R. (2024). Express barcoding with NextGenPCR and MinION for species-level sorting of ecological samples. *Molecular Ecology Resources*, *24*(3), e13922. https://doi.org/10.1111/1755-0998.13922

Voorhuijzen-Harink, M. M., Hagelaar, R., van Dijk, J. P., Prins, T. W., Kok, E. J., & Staats, M. (2019). Toward on-site food authentication using Nanopore sequencing. *Food Chemistry: X*, *2*, 100035. https://doi.org/10.1016/j.fochx.2019.100035

Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348–1365. https://doi.org/10.1038/s41587-021-01108-x

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., … Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, *37*(10), 1155–1162. https://doi.org/10.1038/s41587-019-0217-9

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Wurzbacher, C., Larsson, E., Bengtsson-Palme, J., Van den Wyngaert, S., Svantesson, S., Kristiansson, E., Kagami, M., & Nilsson, R. H. (2019). Introducing ribosomal tandem repeat barcoding for fungi. *Molecular Ecology Resources*, *19*(1), 118–127. https://doi.org/10.1111/1755-0998.12944

Zhu, T., Sato, Y., Sado, T., Miya, M., & Iwasaki, W. (2023). MitoFish, MitoAnnotator, and MiFish pipeline: updates in 10 years. *Molecular Biology and Evolution*, *40*(3). https://doi.org/10.1093/molbev/msad035

1078 Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E.,
1079 Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L.,
1080 Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA metabarcoding-Need for
1081 robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–
1082 1862. https://doi.org/10.1111/mec.15060