



Guidelines for using environmental DNA (eDNA) to monitor the impact of salmon farms on benthic communities

PREPARED BY

JAN PAWLOWSKI, TRISTAN CORDIER, KRISTINA CERMAKOVA

DEPARTMENT OF GENETICS AND EVOLUTION, UNIVERSITY OF GENEVA, SWITZERLAND

ID-GENE ECODIAGNOSTICS™, SWITZERLAND

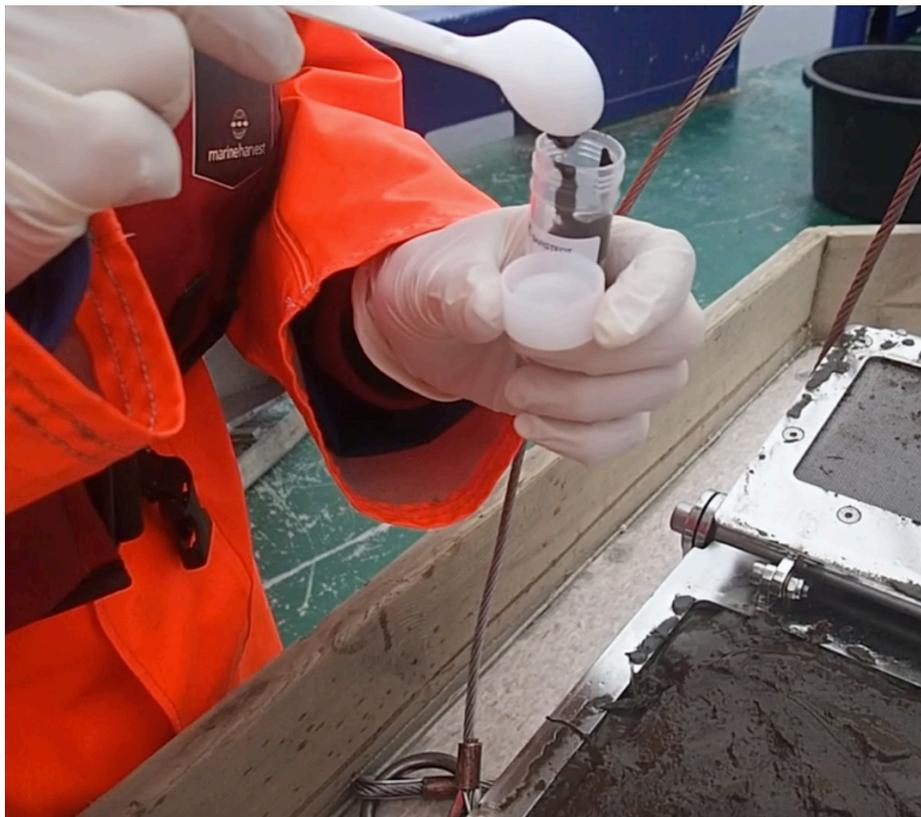


Photo by Tomas Cedhagen

GENEVA, MARCH 29TH, 2019

Table of contents

Introduction	2
General comments on work organization	3
Protocol 1: Sampling sediment DNA	6
Protocol 2: Sediment DNA extraction	9
Protocol 3: Polymerase Chain Reaction (PCR) amplification	11
Protocol 4: High-throughput Sequencing (HTS)	13
Protocol 5: Bioinformatics	14
Acknowledgements	16
References	17

Introduction

Recent development of environmental DNA (eDNA) metabarcoding offers new molecular tools for environmental biomonitoring that use DNA sequences rather than morphological features to identify species and describe biodiversity (Valentini et al. 2016). The high-throughput sequencing technologies allow to obtain millions of eDNA sequences in relatively short time and at low costs. Bioinformatic analysis of these sequences gives global information about species composition of targeted taxonomic groups. This information can be used to assess ecological status through biotic indices (Pawlowski et al. 2018).

There are multiple advantages of using metabarcoding for biomonitoring. Compared to the traditional morphotaxonomic approach, the metabarcoding analyses are:

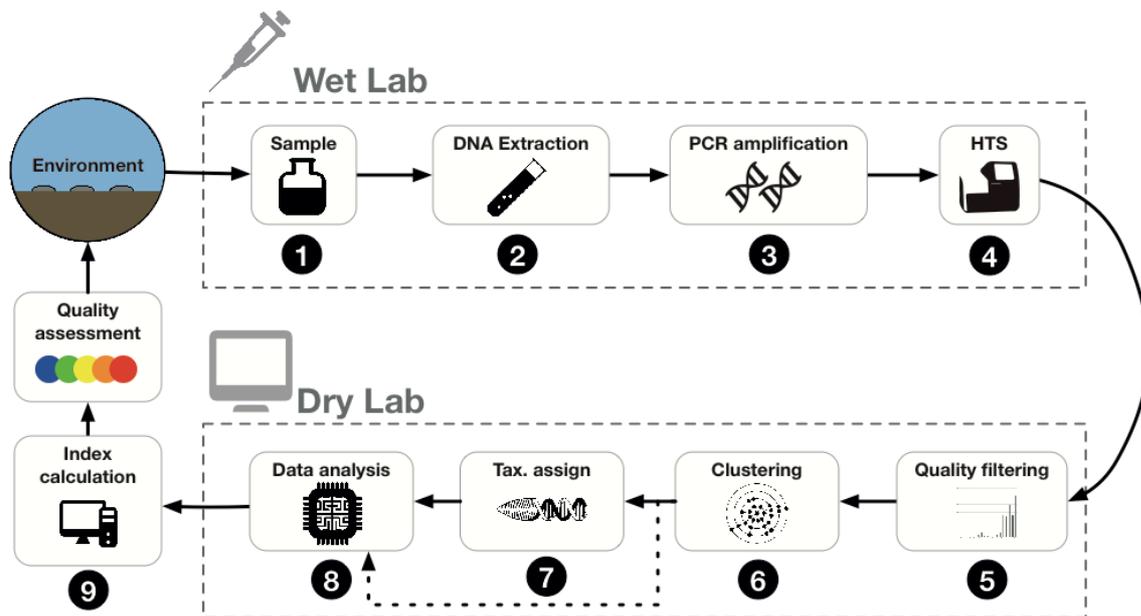
- Faster and cost-effective
- More sensitive – covering larger range of potential bioindicators and allowing accurate identification of small-sized organisms (meiofauna, protists) and morphologically indistinctive cryptic species
- More reliable – based on automatized and standardized protocols of molecular data generation and analysis

The following guidelines describe all steps of metabarcoding approach from DNA sampling to bioinformatic analyses. The guidelines have been developed as a part of the validation study aiming to use eDNA metabarcoding for benthic monitoring of salmon farms in Norway (FHF project 901092). The guidelines refer to several published studies that have been conducted within this project (Pawlowski et al. 2016, Cordier et al. 2017, 2018a, 2018b) as well as to the studies exploring the use of metabarcoding for benthic monitoring of salmon farms in Scotland (Pawlowski et al. 2014, Lejzerowicz et al. 2015, Stoeck et al. 2018a, 2018b, Forster et al. 2018), New Zealand (Dowle et al. 2015, Pochon et al. 2015) and Canada (He et al. 2019).

General comments on work organization

Molecular benthic monitoring is based on eDNA metabarcoding approach, which consists in high-throughput sequencing of amplified target genes (metabarcodes) that are used to analyse the diversity of organisms present in eDNA samples. Its workflow consists of two major parts: molecular (wet lab) part from extracting DNA to sequencing of metabarcodes and bioinformatic (dry lab) part from processing the sequences to analysing the data and assess the environmental quality (Figure 1).

Figure 1. Workflow of eDNA monitoring



Wet Lab The four main steps of molecular part of eDNA monitoring workflow are:

1. Sampling sediment DNA – Few grams of sediment are sampled from the surface of sediment collected from the bottom using a grab or another tool (Protocol 1).
2. Sediment DNA extraction – The molecules of DNA (extracellular DNA and DNA of living organisms, their tissue fragments and organelles) contained in the sediment are isolated (Protocol 2).
3. Polymerase Chain Reaction (PCR) amplification – During this step millions of copies of targeted genomic region are amplified in a series of enzymatic reactions (Protocol 3).
4. High-throughput Sequencing (HTS) – The PCR-amplified metabarcodes are used as template for massively parallel DNA sequencing producing millions of metabarcoding sequences (Protocol 4)

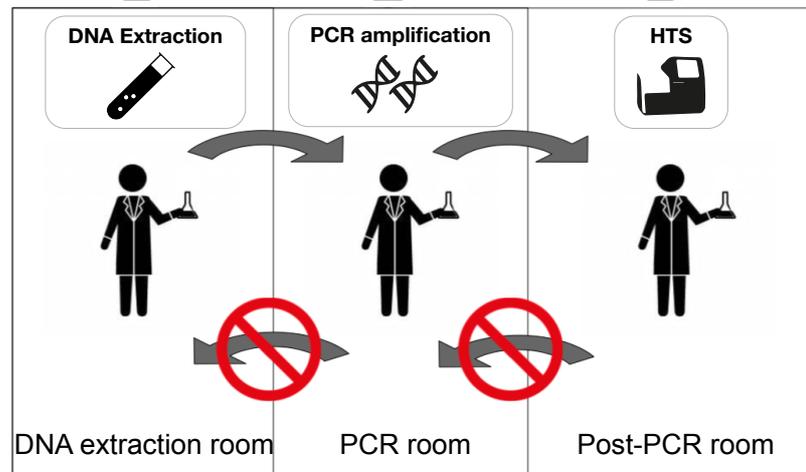


Dry Lab Bioinformatic part of eDNA monitoring workflow comprises:

5. Quality-filtering – Metabarcoding sequences with a low quality and/or ambiguous bases are removed. The paired-end sequences are merged into a contiguous full-length sequence and potential chimeras are removed.
6. Clustering – High-quality sequences are clustered according to their similarity to one another and grouped into operational taxonomic units (OTUs).
7. Taxonomic assignment – OTUs with sequence found in a reference database are assigned to taxa.
8. Data analysis – The list of OTUs serves to analyse the taxonomic composition of each sample and their relation to environmental variables.
9. Index calculation – Biotic indices used to assess the ecological status of the samples are either calculated from the assigned OTUs or predicted from the whole dataset using machine learning approach.

Organization of the wet lab for eDNA monitoring should comply with special facility design and workflow guidelines. These constraints are time and cost demanding but they are essential to get trustable data for further analysis. Due to the high sensitivity of PCR technique, a laboratory performing eDNA analysis should be divided into at least three physically separate rooms (Figure 2).

Figure 2. Wet lab facility and work organization



The separation of workspace is based on DNA concentration gradient:

- DNA extraction room (low DNA concentration)
- PCR room (DNA amplification)
- Post-PCR room (high DNA concentration)

Useful tips to prevent contamination as well as cross-contamination between samples:

- ☑ In each room, use gloves and a lab coat dedicated for that room.
- ☑ Regularly clean and DNA-decontaminate the workspace and items.
- ☑ Follow a unidirectional workflow from low to high DNA concentration to mitigate contamination with amplified DNA from previous assays.
- ☑ No material and device transit against the DNA concentration gradient.
- ☑ No work is performed against the DNA concentration gradient within the same day. For example, a work session in the Post-PCR room can never be followed by a work session in the PCR room or the DNA extraction room.

To ensure reproducibility, follow consistently the established protocols and keep track of the samples and all the processing steps.

Protocol 1: Sampling sediment DNA

We recommend using the Life Guard™ (Qiagen) solution for preserving sedimentary DNA. Its properties allow maintaining the integrity of DNA by preventing DNase activity. In consequence, the immediate freezing of samples is not mandatory. Using Life Guard™ facilitates sampling in remote areas where freezing facilities are not available and it allows to transport samples without freezing.



- ☑ Store and handle all the equipment and items necessary for sampling in a separate room kept apart from PCR-amplified DNA or other highly concentrated DNA sources.



Sampling equipment:

- sterile latex powder free gloves
- sterile spoons 2,5ml (e.g. VWR, ref. 231-0513)
- 20ml tubes (e.g. Sarstedt, ref. L0262-7)
- Life Guard™ solution (QIAGEN)
- Zip lock slider bags
- stickers to identify the station
- additionally:
- cooler bricks or ice packs
- portable cooler for transportation

Prior to sampling:

- Put a sticker on each zip lock slider bag
- Fill each 20ml tube with 5ml of Life Guard™
- Prepare 1 **kit per station** inside a zip lock slider bag labelled with a sticker containing:
 - 1 pair of sterile gloves
 - 2 sterile spoons (1 per grab)
 - 4 tubes (2 per grab) each containing 5ml of Life Guard™
- Freeze the cooler bricks or ice packs

Collecting sediment DNA samples

The method was developed for soft sediments; localities/grabs containing mostly shelly sand and gravel should be avoided.

The samples for eDNA analysis are taken from 2 grabs per station. They can be taken after the measurement of redox, and before the other subsampling or sieving.

If a macrofauna study is also to be done, the sediment sample for eDNA analysis must be taken from the same grab that is sieved for macrofauna study.



- ☑ Avoid:
 - grabs with very small volume of sediment
 - grabs containing mostly shelly sand or gravel.
- ☑ Wear gloves when collecting the samples.
- ☑ Take the samples for eDNA analysis rapidly once the grab is on board.

1. Gently bring on board the Van Veen grab or another tool used to sample the sediment, try to not agitate the sediment and keep its surface intact.
2. If the grab contains water, suck it out in order to access the sediment surface.
3. Put the gloves on.
4. Scoop about 5 to 10ml of sediment from the surface top 2cm into the tube with the sterile spoon probing the whole grab surface rather than sampling single spot. Make sure the final volume is between 10 and 15ml.
5. Try to not spill the Life Guard™ solution and close the tubes firmly.
6. Shake the tubes to mix thoroughly the sediment and the solution.
7. Take 2 samples per grab. Change spoons between grabs.
8. Label tubes sequentially according to following model:
 - Site name, Station 1: grab 1 tube 1, 2; grab 2 tube 3, 4*
 - Site name, Station 2: grab 1 tube 5, 6; grab 2 tube 7, 8..*
9. Place the tubes in plastic bags (one bag per station), fill the labels on bag (Site, Station, Date).
10. Store the tubes in a cooler after collecting.

At the end of the day, place the samples in a freezer at -20°C. If it is not possible, keep them in a refrigerator or a cooler.

Storage and shipping

Transport the samples in a cooler or a polystyrene box with cooler bricks or ice packs. The Life Guard™ solution present in the tubes with the samples allows to preserve the eDNA during the transport even when the samples happen to be exposed to ambient temperature for several hours. The sediment samples in Life Guard™ solution can be stored at -20°C for several weeks.

Protocol 2: Sediment DNA extraction

PowerSoil® DNA extraction kits (QIAGEN) provide a good combination of DNA yield, purity and diversity coverage (Lekang et al., 2015). PowerSoil® DNA kits allowing an extraction from up to 0,25g of sediment are suitable for surveys of microbial communities (e.g. standard for the Earth Microbiome Project). Here, we recommend using DNeasy® PowerMax® Soil Kit (QIAGEN) that allows processing up to 10g of sediment and gives a better diversity coverage.



Photo: www.qiagen.com

Taking several replicates contributes to reduce the effects of micro-patchiness in distribution of benthic species (Lejzerowicz et al. 2014). Extraction can be done either from each replicate or from a homogenised pool of replicates for each grab.



- ☑ Wear gloves and lab coat reserved only to this room.
- ☑ Store and use all the equipment and items necessary for DNA extraction in a dedicated DNA extraction room.
- ☑ Avoid circulation of items from the DNA extraction room.
- ☑ Never perform extractions following a work session in the PCR or the Post-PCR room.
- ☑ Use filter tips.
- ☑ The extraction room may be equipped with a fume hood to limit exposure of the personnel to the extraction solutions.
- ☑ The hood may be equipped with a UV-lamp to sterilize the workspace and the items between working sessions.

Before you start:

- Clean up all instruments, bench tops and racks surface that will be used during the DNA extraction with DNA Away™ cleaning solution (Thermo Scientific) or equivalent.
- Thaw the frozen sediment samples at room temperature and immediately put them on ice or in refrigerator.
- Prepare cold molecular grade ethanol (10,5ml per sample) on ice or in refrigerator.

1. Add about 5 to 7g of wet sediment to the PowerBead Tubes and centrifuge at 2500g for 5min. Discard the supernatant of Life Guard™ preservation solution.
2. Add 15ml of PowerBead solution to the PowerBead Tubes containing sediment and vortex vigorously.
3. Add 1,2ml of solution C1 and vortex.
4. Place the tubes horizontally in the vortex adapter and vortex for 10min at the highest speed (point 4 in the Quick-Start protocol of DNeasy® PowerMax® Soil Kit).
5. Continue according to the instructions in DNeasy® PowerMax® Soil Kit manual.

The final 5ml of eluted DNA may be concentrated by ethanol precipitation:

1. Add 200µl of 5M NaCl and 10,4ml of cold ethanol to the eluted DNA, invert to mix and incubate at -20°C for a minimum of 30min (can be done over-night).
2. Centrifuge at 2500g for 30min and discard the supernatant. Attention to not lose the pellet of precipitated DNA.
3. Let the residual ethanol evaporate at ambient air (avoid over-drying).
4. Resuspend the precipitated DNA with 400µl of sterile 10mM Tris solution.
5. Prepare an aliquot of each DNA for further application.

The DNA should be preserved at -20°C.

Protocol 3: Polymerase Chain Reaction (PCR) amplification

PCR primers

The choice of barcoding gene depends on taxonomic group. Because different genomic regions evolve at different rates it is crucial to select the region that provides the best taxonomic resolution for a given group. Here, we recommend using 37F region of nuclear 18S rRNA gene for benthic foraminifera (Pawlowski et al., 2014), mitochondrial cytochrome oxidase 1 gene COI (the BOLD standard), targeting principally Metazoa, and the V1V2 and V4 hypervariable regions of nuclear 18S rRNA gene (Fonseca et al. 2010, Stoeck et al. 2010), targeting all eukaryotes.

18S V1V2 is recognized among the best markers for eukaryotic diversity and amplifies a high proportion of metazoan taxa. The paper of Cordier et al. 2018a shows that this marker is suitable for predicting biotic indices.

The COI marker is recommended to be used in parallel with 18S V1V2 to obtain a complementary view of macrofaunal community. Although a high proportion of the COI sequences obtained from marine sediment DNA are impossible to identify, the COI reference database contains more macro-invertebrate species and provides a better taxonomic resolution than 18S V1V2.

The sequences of primers used for amplification of each marker are indicated in Table 1.

Table 1. Primers sequences

Taxonomic group	Gene region	Primer name	Sequence (5' - 3')
Foraminifera	37F	s14F1	AAGGGCACCACAAGAACGC
		s15	CCACCTATCACAYAATCATG
Metazoa	COI	mICOLintF	GGWACWGGWTGAACWGTWTAYCCYCC
		dgHCO2198	TAAACTTCAGGGTGACCAAARAAYCA
Eukaryota	V1V2	F04mod	GCTTGWCTCAAAGATTAAGCC
		R22mod	CCTGCTGCCTTCCTTRGA
Eukaryota	V4	TAReukFWD1	CCAGCASCYGCGGTAATTCC
		TAReukREV3	ACTTTCGTTCTTGATYRA

Using a unique combination of tagged primers bearing 8 nucleotides attached at 5'- extremity for each PCR-amplified sample enables the multiplexing of PCR products in sequencing libraries (Pawlowski et al. 2014, Esling et al. 2015).

PCR amplification



- ☑ Wear gloves and lab coat reserved only to this room.
- ☑ Store and use all the equipment and items necessary for PCR amplifications in a room dedicated exclusively to PCR amplifications.
- ☑ Avoid circulation of items from the PCR room.
- ☑ Never perform PCR amplifications following a work session in the Post-PCR room.
- ☑ Use filter tips.
- ☑ Include PCR negative controls (for each unique combination of tagged primers) in each session to verify that no contamination occurred.
- ☑ In the PCR room, use aliquots of eDNA extracts, prepared in the DNA extraction room, to avoid unspecific introduction of PCR components into the stock eDNA.
- ☑ Work in a UV-sterilized PCR cabinet with no air circulation to prevent cross-contaminations.

PCR reactions should be performed in three replicates for each eDNA sample.

We recommend amplifying 1µl of eDNA in 25µl of reaction volume using FastStart™ Taq polymerase (Sigma-Aldrich) with PCR Reaction Buffer containing 20mM MgCl₂, adding Bovine Serum Albumin at 200µg/ml.

Assemble all the components of the reaction in a PCR cabinet in the PCR room.

Use a thermocycler placed in the Post-PCR room to amplify the target gene region using the conditions described in Table 2.

Table 2. Conditions for PCR amplification of each target gene region

37F: Pre-incubation at 95°C for 5min

- 40 cycles composed of:
 - 94°C for 20sec
 - 52°C for 20sec
 - 72°C for 20sec
- Incubation at 72°C for 2min.

V1V2: Pre-incubation at 95°C for 5min

- 35 cycles composed of:
 - 95°C for 30sec
 - 57°C for 30sec
 - 72°C for 1min
- Incubation at 72°C for 5min.

COI: Pre-incubation at 95°C for 5min

- 16 cycles composed of:
 - 94°C for 10sec
 - 62-47°C for 30sec
(1°C decrease every cycle)
 - 72°C for 1min
- 35 cycles composed of:
 - 95°C for 10sec
 - 46°C for 30sec
 - 72°C for 1min
- Incubation at 72°C for 2min.

V4: Pre-incubation at 95°C for 5min

- 10 cycles composed of:
 - 94°C for 20sec
 - 57°C for 45sec
 - 72°C for 1min
- 35 cycles composed of:
 - 94°C for 20sec
 - 47°C for 45sec
 - 72°C for 1min
- Incubation at 72°C for 2 min.

Protocol 4: High-throughput Sequencing (HTS)

The performance of different high-throughput sequencing platforms is compared in Quail et al. 2012, Frey et al. 2014 and Ren et al. 2015. Here, we recommend using Illumina MiSeq technology because it provides the best trade-off for amplicon sequencing with the sequences length up to 600 nucleotides and a maximal output of 24 million sequences.



- ☑ Wear gloves.
- ☑ Perform all the manipulations downstream of the PCR amplification step in a specific room for amplified DNA.
- ☑ Avoid circulation of items from the Post-PCR room.
- ☑ Avoid working in DNA extraction room and PCR room after a work session in the Post-PCR room.
- ☑ Use filter tips.

1. Combine the three replicate PCR products for each sample.

If the PCR products are tagged:

- Quantify the combined replicates for each sample. Using high-resolution capillary electrophoresis (e.g. QIAxcel System, QIAGEN) is recommended.
 - Pool equimolarly between 500 and 1000ng of amplicons for each multiplexed library. We recommend to pool a maximum of 50 differently tagged samples.
2. Purify the PCR products (e.g. using High Pure PCR Product Purification kit, Roche)
 3. Quantify the purified PCR products using fluorometric method (e.g. Qubit HS dsDNA kit, Invitrogen).
 4. Depending on the amount of DNA, prepare the libraries using TruSeq® Nano DNA Sample Preparation Kit (100-200ng of DNA for a library) or TruSeq® DNA PCR-Free Library Preparation Kit (700-1000ng of DNA for a library). Follow the manufacturer's instructions, without performing the initial DNA fragmentation step.
 5. Quantify the libraries by qPCR using Kapa Library Quantification Kit for Illumina Platforms (Kapa Biosystems). 1:10'000 and 1:20'000 independent (not serial) dilutions are recommended using at least 2ul of each library and the Standard 0. Perform all the reactions in triplicates.
 6. Sequence on a MiSeq instrument (Illumina) using paired-end sequencing with 502 cycles. The sequencing depth, i.e. the number of sequences per sample, should average at least 125'000 sequences, meaning that a maximum of 120 samples can be sequenced using MiSeq Reagent Kit v2 and a maximum of 192 samples using MiSeq Reagent Kit v3.

Protocol 5: Bioinformatics

The obtained HTS data are analysed using well-established and publicly available algorithms and software. SLIM, a web application described in Dufresne et al., 2019, allows to create and reproducibly execute the processing pipeline through a graphic user interface. The stringent processing parameters that we recommend here aim to ensure keeping only high-quality data for the analysis. For reproducibility, keep track of your processing pipeline and all the processing parameters.

Quality-filtering



1. Raw sequencing data can be quality-filtered by removing sequences with a low quality (mean quality score below 30).
2. For multiplexed libraries, demultiplex the sequences allowing no mismatch in the tagged primers.
3. Merge paired-end reads by aligning them into a contiguous long-length sequence. We recommend the VSEARCH toolkit (Rognes et al., 2016), with a minimum overlap of 40 base pairs and five mismatches allowed.
4. Remove potential chimeras. We recommend the UCHIME de novo algorithm (Edgar, 2010) with the default settings.
5. Remove all sequences with ambiguous bases.

Clustering



1. Cluster the sequences into Operational Taxonomic Units (OTU). We recommend using the SWARM algorithm (Mahé et al., 2015) with the default resolution ($d = 1$).
2. Discard rare OTUs (we recommend below 100 reads).

Taxonomic assignment



Compare the OTUs representative sequences against a reference sequence database for taxonomic assignments. We recommend using the QIIME toolkit (Caporaso et al., 2010) with 95% minimum similarity and default parameters (uclust method, Edgar, 2010). We recommend SILVA database (Quast et al., 2013) for the nuclear ribosomal 18S marker and MIDORI database (Machida et al., 2017) for the COI marker.

Data analysis



The downstream data analyses can be performed in the R programming environment (R Core Team).

- We recommend discarding the samples with less than 10'000 reads.

Index calculation



Assess the ecological status of the samples.

- Compute the Benthic Biotic Indices with metazoan assigned OTUs. We recommend using the BBI R package (Cordier and Pawlowski, 2018) that allows to compute AMBI, NSI, ITI, ISI and Bentix values and the normalised Ecological Quality Ratio index (nEQR).

and/or

- Predict the biotic indices. Machine learning approach described in Cordier et al., 2017, Cordier et al., 2018a allows to predict the values of AMBI, NSI, NQI 1, ISI, as well as Shannon and nEQR. The predictive model was trained on metabarcoding data obtained for salmon farming sites in Norway, for which the ecological status was known (training datasets available at the Sequence Read Archive under BioProject PRJNA376130 for foraminifera 37F and PRJEB23641 for eukaryotes V1V2 and V4). The biotic indices values are predicted for each sample and the values are averaged into a final predicted value for each station.

Acknowledgements

The protocols presented in these guidelines were developed in the framework of the project 901092 funded by The Norwegian Seafood Research Fund (FHF).

The authors thank Laure Apothéloz-Perret-Gentil, Inès Barrenechea, Tomas Cedhagen, Yoann Dufresne, Philippe Esling, Anders Lanzén, Franck Lejzerowicz, Joana Visco and Sofia Wyler for participation in this project and help in development of different steps described in these guidelines.

References

- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, ... Knight R (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7(5), 335–336.
- Cordier T, Esling P, Lejzerowicz F, Visco J, Ouadahi A, Martins C, Cedhagen T, Pawlowski J. (2017) Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning. *Environ Sci Technol*. 51(16):9118-9126.
- Cordier T, Forster D, Dufresne Y, Martins CIM, Stoeck T, Pawlowski J (2018a) Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Mol Ecol Res* 18(6):1381-1391.
- Cordier T, Lanzén A, Apothéoz-Perret-Gentil L, Stoeck T, Pawlowski J (2018b) Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. *Trends Microbiol*. S0966-842X(18)30238-5.
- Cordier T & Pawlowski J (2018) BBI: an R package for the computation of Benthic Biotic Indices from composition data. *Metabarcoding and Metagenomics*. 2. e25649. 10.3897/mbmg.2.25649.
- Dowle, E., Pochon, X., Keeley, N., Wood, S., 2015. Assessing the effects of salmon farming seabed enrichment using bacterial community diversity and high-throughput sequencing. *FEMS Microbiol. Ecol*. 91. <https://doi.org/10.1093/femsec/fiv089> (fiv089).
- Dufresne Y, Lejzerowicz F, Perret-Gentil LA, Pawlowski J, Cordier T (2019) SLIM: a flexible web application for the reproducible processing of environmental DNA metabarcoding data. *BMC Bioinformatics* 20(1):88. 10.1186/s12859-019-2663-2. 10.1186/s12859-019-2663-2. 21-02-2019.
- Edgar RC, Haas BJ, Clemente JC, Quince C, & Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 27(16), 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>.
- Esling P, Lejzerowicz F, & Pawlowski J (2015) Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic acids research* 43(5), 2513-24.
- Fonseca VG, Carvalho GR, Sung W, Johnson HF, Power DM, Neill SP, ... Creer S (2010) Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature Communications*, 1(7), 98.
- Forster D, Filker S, Kochems R, Breiner H-W, Cordier T, Pawlowski J, Stoeck T. (2018) A comparison of different ciliate DNA metabarcode genes as bioindicators for environmental impact assessments of salmon aquaculture. *J. Eukaryotic Microbiol*. <https://doi.org/10.1111/jeu.12670>
- Frey KG, Herrera-Galeano JE, Redden CL, Luu TV, Servetas SL, Mateczun AJ, Mokashi VP, Bishop-Lilly KA (2014) Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC Genomics*. 4;15:96.
- He X, Sutherland TF, Pawlowski J, Abbott CL (2019) Responses of foraminifera communities to aquaculture-derived organic enrichment as revealed by environmental DNA metabarcoding. *Mol Ecol*. 2019 Jan 7. doi: 10.1111/mec.15007. 07-01-2019.
- Lejzerowicz F, Esling P, Pillet LL, Wilding T a., Black KD, Pawlowski J (2015) High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. *Sci Rep* 5:13932.
- Lejzerowicz, F, Esling P, Pawlowski J. (2014) Patchiness of deep-sea benthic Foraminifera across the Southern Ocean: insights from high-throughput DNA sequencing. *Deep-Sea Research II*, 108:17-26.
- Lekang K, Thompson EM, Troedsson C (2015) A comparison of DNA extraction methods for biodiversity studies of eukaryotes in marine sediments. *Aquatic Microbial Ecol*. 75:15-25.
- Machida RJ, Leray M, Ho SL, Knowlton N (2017) Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Sci. Data* 4, 1–7. <https://doi.org/10.1038/sdata.2017.27>.

- Mahé F, Rognes T, Quince C, De Vargas C, & Dunthorn M (2015) Swarm v2: Highly-scalable and high-resolution amplicon clustering. *PeerJ* 3 e1420 <https://doi.org/10.7717/peerj.1420>.
- Pawlowski J, Esling P, Lejzerowicz F, Cedhagen T, Wilding TA (2014) Environmental monitoring through protist next-generation sequencing metabarcoding: Assessing the impact of fish farming on benthic foraminifera communities. *Mol Ecol Resour* 14:1129–1140.
- Pawlowski J, Esling P, Lejzerowicz F, Cordier T, Visco JA, Martins CIM, Kvalvik A, Staven K, Cedhagen T (2016) Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. *Aquac Environ Interact* 8:371–386.
- Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, ... Kahlert M (2018) The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total Environ.* 637-638():1295-1310.
- Pochon X, Wood SA, Keeley NB, Lejzerowicz F, Esling P, Drew J, Pawlowski J (2015) Accurate assessment of the impact of salmon farming on benthic sediment enrichment using foraminiferal metabarcoding. *Mar Pollut Bull.* 100(1):370-82.
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, 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.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41. <https://doi.org/10.1093/nar/gks1219>.
- Ren X, Hu Y, Yang F, Xiao Y, Li L, Dong J, . . . Jin Q (2015) Clinical utility comparison of two benchtop deep sequencing instruments for rapid diagnosis of newly emergent influenza infections. *Clinical Microbiology and Infection.* 21 (3), 290–e1.
- Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, Richards TA (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19, 21–31. doi: 10.1111/j.1365-294X.2009.04480.x
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F, 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4 <https://doi.org/10.7717/peerj.2584>. e2584.
- Stoeck T, Kochems R, Forster D, Lejzerowicz F, Pawlowski J. (2018a) Metabarcoding of benthic ciliate communities shows high potential for environmental monitoring in salmon aquaculture. *Ecol. Indicators* 85:153-164
- Stoeck T, Frühe L, Forster D, Cordier T, Martins CIM, Pawlowski J. (2018b) Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic footprint of salmon aquaculture. *Marine Pollution Bulletin* 127:139-149
- The Barcode of Life Data System (BOLD) (<http://www.barcodinglife.org>).
- The Earth Microbiome Project – DNA extraction protocol ([dx.doi.org/10.17504/protocols.io.nutdwn](https://doi.org/10.17504/protocols.io.nutdwn)).
- Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, ... Dejean T (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.* 25:929–942.