Marine Finfish Farms

Guidance on the use of DNA in monitoring compliance with seabed mixing zone area limits

February 2023

Version 1.0

Every day SEPA works to protect and enhance Scotland's environment, helping communities and businesses thrive within the resources of our planet.

We call this **One Planet Prosperity**

Introduction

This document provides guidance to operators of marine finfish farms on using DNA analyses to monitor compliance with the seabed mixing zone boundary biological standard specified in their farms' permits.

DNA data collected and analysed by, or on behalf of, finfish farm operators in accordance with this guidance will be accepted by SEPA for compliance assessment purposes in lieu of data on seabed invertebrate animals.

Requirements for the collection and analysis of the latter are set out in SEPA's soft-substrate sampling and biological testing performance standards under its quality assurance certification scheme (MACS)[1].

This guidance describes how samples need to be collected, preserved, and analysed to generate suitable DNA data. The methods listed below represent SEPA's preferred approach at time of writing to allow for as standardised an approach as practicable. However, once DNA analysis has become more established, there will be scope to include other methodologies providing the same standard of data quality and reproducibility can be demonstrated.

Process overview

Figure 1 outlines how SEPA will take account of sediment bacterial DNA analysis results submitted by marine finfish farm operators when assessing compliance with seabed mixing zone area limits.

---

[1] https://www.sepa.org.uk/regulations/how-we-regulate/operator-monitoring/
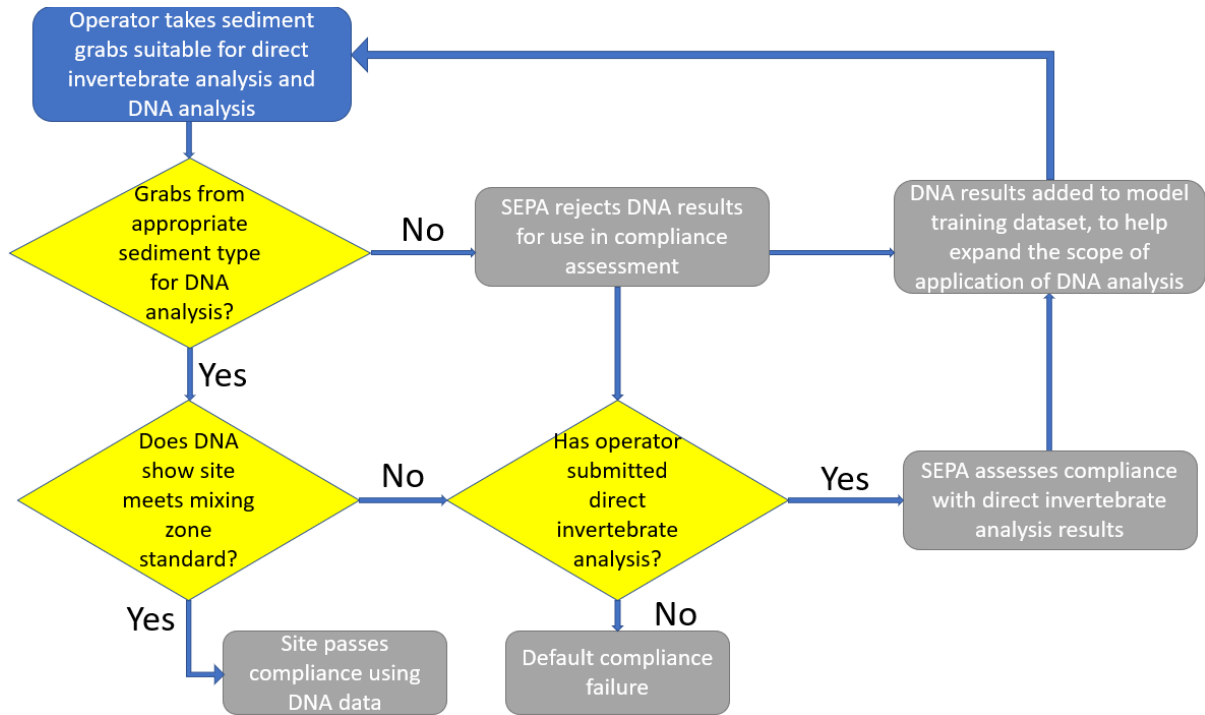
Figure 1. Use of DNA analysis results in assessing compliance with seabed mixing zone area standard

Operators may provide DNA results to SEPA for all seabed monitoring stations intended for mixing zone limit compliance assessment or a combination of DNA results for some stations and seabed invertebrate analysis results for others.

If they are suitable, SEPA will use the reported DNA results to derive infaunal quality index (IQI) scores for the monitoring stations.

The method for generating IQI scores from DNA analyses was developed by a collaborative research project managed by the Sustainable Aquaculture Innovation Centre, with the Scottish Association for Marine Science as the lead academic research partner. It is based on a machine learning algorithm trained using a dataset of matched bacterial DNA and invertebrate sample results.

The seabed samples used in training the algorithm encompassed a range of sediment types. IQI scores can be generated for compliance assessment purposes from DNA results for samples from any seabed types within the range used to train the algorithm (Annex 1).

The range of sediment types covered by the algorithm will be expanded as and when sufficient data from additional seabed types is available to enable the algorithm to be re-trained. Any relevant and suitable matched DNA analysis and invertebrate analysis results submitted by operators will be used for this purpose along with results from surveys undertaken by SEPA.

Monitoring station IQI scores, whether derived from suitable DNA analysis results or from analyses of seabed invertebrates, will be used in assessing compliance with the farm's mixing zone area limit.

The IQI scores generated by the algorithm from DNA results have a bias towards being lower than those calculated from invertebrate analyses. Consequently, the estimated distance from a farm along a monitoring transect to the IQI score indicating good seabed quality may be greater when using DNA-derived IQI scores than when using IQI scores calculated from invertebrate analyses.

For those farms with areas of deposition on the seabed close to the limit of the farms' permitted mixing zone areas, operators should plan to demonstrate compliance based on IQI scores calculated from invertebrate analyses. For farms operating well within their mixing zone area limits, the use of DNA-derived IQI scores to demonstrate compliance is expected to significantly reduce the time needed for sample analysis, enable faster assessments of compliance and offer potentially substantial efficiency savings.

3

The generation of IQI scores from DNA analysis results and assessments of compliance with the mixing zone area limits will be undertaken by SEPA and the results fed back to the operator. Subject to the required data being submitted in the correct format and without errors, SEPA will aim to feedback assessment results within 4 weeks.

Information about how IQI scores are generated and assessments of compliance with mixing zone area limits are described in the SEPA Seabed mixing zone limit: compliance assessment methodology[2].

In due course, SEPA is aiming to develop a web-based service that will enable operators to use an on-line application to generate IQI scores and to check the area of their farms' mixing zones in real time.

The following sections describe how samples should be:

- Collected for DNA analysis.
- Preserved.
- Analysed to generate suitable DNA information.

1. DNA sample collection

1.1 DNA samples should be taken directly from the same grabs used to collect samples of benthic invertebrates using the method specified in SEPA's MACS performance standard for sampling of soft substrate[3]. The samples should be taken from the sediment surface of the grabs prior to the grabs being emptied and their contents sieved.

---

[2] https://www.sepa.org.uk/media/594223/seabed-mixing-zone-limit_compliance-assessment-method.pdf
[3] https://www.sepa.org.uk/media/594190/macs-ffa-ps-01.pdf

4

1.2   Grabs with a surface area bite size of 0.1 m^2 or 0.045 m^2 must be utilised.

1.3   The mouth of the grab should be carefully opened slightly to let water drain out and expose the sediment. If the sediment surface looks like it has been disrupted, for example by pressing against the doors of the grab, the grab should not be used for DNA sampling and a new grab should be taken. In soft muds, over-biting of the grab can be reduced by removing weights.

1.4   Samples should be taken across the surface of the grab as illustrated in Figure 2. When using a 0.1 m^2 grab, 6 scoops of approximately 2-3 ml per scoop should be taken from each point in the grab and combined into one sample in a 50 ml falcon tube. When using a 0.045 m^2 grab, 2 grabs are required, with 3 scoops from each grab taken and combined into one sample. Care should be taken not to take samples too deeply (scoops should be from no greater than 1 cm depth). As a minimum, a final volume of 10 ml per sample should be collected.
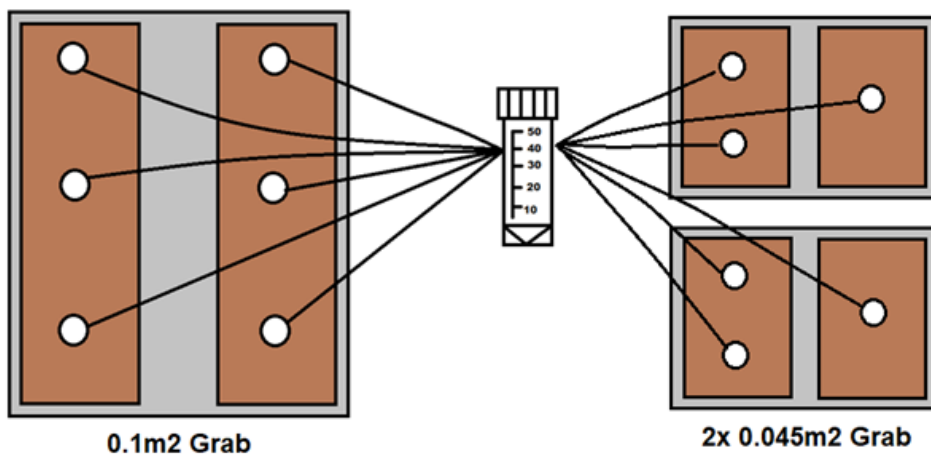
Figure 2. Example of points in a 0.1 m^2 or 2x 0.045 m^2 grabs from which to take sediment scoops, combined in a single sample tube.

1.5   After the DNA sample has been taken from the grab, the remainder of the sample in the grab must be processed and the organisms within it preserved in

5

accordance with the MACS performance standard[4]. The preserved sample must be stored for later invertebrate analysis, should such analysis be required to demonstrate or corroborate compliance.

1.6   The PSA (particle size analysis) sample for the monitoring station should be used to provide PSA data for subsequent derivation of the IQI score for the DNA sample as well as for the calculation of the IQI score from invertebrate sample analysis, if undertaken.

1.7   The meta-data required to accompany DNA samples is largely the same as for benthic invertebrate samples:
- GPS coordinates
- Site name, transect number and sample station number
- Time and date
- Water depth
- Grab size used
- Sediment volume relative to grab volume (as a fraction i.e., ¼, ½, ¾ etc.)
- Sediment descriptor

With the addition of:
- Time of preservation (i.e. the time a sample is first frozen)
- Approximate volume per DNA sample (ml)
- Additional comments

2.   DNA sample preservation

2.1   Once a DNA sample has been taken, the sample must be labelled, sealed by screwing the lid on tightly and placed into a sealable bag and frozen if possible.

---

[4] https://www.sepa.org.uk/media/594190/macs-ffa-ps-01.pdf

2.2   If no freezer is available during surveying or transport from the vessel, the sample must be stored on ice in an insulated cooler box/bag.

2.3   A clean spatula should be used for each sample. Spatulas can be decontaminated and reused using the following procedure; soak spatulas in 10% bleach (prepared on day of use) for 10 minutes, then soak for a further 5 minutes in 5% Lipsol before triple rinsing in copious volumes of fresh deionised/distilled water.

NB Bleach can negatively impact molecular methods, which could affect results, so every care should be taken to ensure that decontaminated equipment is thoroughly rinsed before it is reused.

2.4   Samples must be chilled and transferred to a freezer (-20°C or below) for storage as soon as possible (ideally within 4 hrs of collection). The time of preservation is recorded as the time samples are first frozen and not when placed on ice as an interim cooling measure.

2.5   To preserve sample integrity, the number of freeze-thaw cycles that a sample is subjected to prior to DNA extraction should be kept to an absolute minimum and avoided altogether where practicable.

3. DNA sample analysis

3.1   The procedures for DNA sample analysis described in this section should be followed to ensure consistency in the application of laboratory processes. Any deviation(s) must be reported when data is submitted to SEPA. All DNA results must be robustly assessed for both quality and quantity.

3.2   Laboratories performing this analysis should be set up to ensure no cross-contamination can occur by separating pre and post PCR-amplification tasks. At a minimum, all pre-PCR work should take place in a dedicated space. The analysis steps are summarised in Figure 3.

3.3   DNA extraction should be undertaken following the manufacturer's instructions and using a DNeasy PowerSoil Kit or DNeasy PowerMax Soil Kits (Qiagen)[5]. In brief, after a physico-chemical lysis stage microbial DNA is purified and collected via spin column technology.
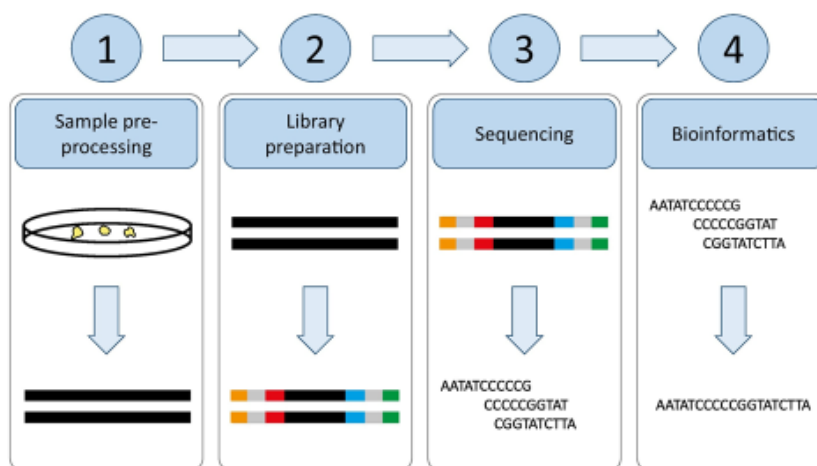

Figure 3. An illustration of the laboratory steps included in this procedure.

3.4   The washes and buffers supplied in commercial kits should be used to help eliminate PCR inhibitors, such as humic substances, that may naturally occur in environmental samples.

---

[5] https://www.qiagen.com/us/resources/resourcedetail?id=9bb59b74-e493-4aeb-b6c1-f660852e8d97&lang=en
[5] https://www.qiagen.com/gb/resources/resourcedetail?id=8ba278a1-2b94-4d79-942b-19310489f09f&lang=en

3.5   An extraction blank must be run for every batch of extractions. Extraction blanks should use the same weight-by-volume of purified water in place of sediment.

3.6   The extraction blanks should be processed downstream in an identical fashion to DNA samples.

3.7   The bacterial gene marker-fragment-primer combination shown in Table 1 must be used to generate sequence output. For the primer sequences, degenerate nucleotide bases are underlined.

Table 1: Marker indicating the specific region, forward (F) and reverse (R) primers, approximate amplicon length (AL, in base pairs) and reference.

Marker	Primer name	Sequence	AL	Reference
16s (V3/V4)	16s_V3-V4_F	CCTACGGGNGGCWGCAG (17 bp)	~460	(Klindworth, Pruesse et al. 2013)[6]
	16s_V3-V4_R	GACTACHVGGGTATCTAATCC (21 bp)		

3.8   The V3/V4 region of the 16S marker has been specifically incorporated into a standard protocol by Illumina[2] for use on their high-throughput sequencers, and this should be followed.

---

[6] Klindworth A., Pruesse, E., Schweer T., Peplies J., Quast C., Horn M. & Glockner F. O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies; Nucleic Acids Research, Volume 41, Issue 1, https://doi.org/10.1093/nar/gks808

3.9   The workflow follows a two-step polymerase chain reaction (PCR) that includes a primary PCR amplification (Amplicon PCR), in triplicate (pooled prior to bead clean-up), where the selected marker fragment is amplified using specific primers (Table 1), with Illumina overhang adapters attached to the 5' end of the primers.

3.10   A second PCR (index PCR) attaches Illumina-specific sequencing adapters and indices to the first PCR product according to the dual indexing strategy. A combination of forward and reverse indices allows for samples to be uniquely identified by the combination of these tags. This allows for demultiplexing of samples after sequencing to link sequences and samples. Indexed libraries should then be quantified and normalised before being combined into a single pool.

3.11   Prior to pooling it is recommended that libraries are qPCR or fluorometrically quantified and assessed for size (final target ~630bp).

3.12   Before sequencing, the final pooled library should be denatured, diluted to ensure adequate clustering and have a PhiX internal control added. The final amplicon library, which should contain at least 5% PhiX, should then be loaded onto a Miseq reagent cartridge capable of 600 cycle runs. NB higher PhiX concentrations may be necessary to mimic diversity and ensure adequate sequencing.

3.13   To reduce bias, the following must be used for PCR reactions:

(i) 2x KAPA HiFi HotStart ReadyMix (default recommended option)
(ii) bead-based purification systems providing they selectively bind fragments of the appropriate size.
(iii) different dual-indexed adapters provided they don't increase the likelihood of crosstalk and misassigned reads. NB heterogeneity spacers may be used

10

to increase library diversity and the number of high-quality reads generated per run.

3.14   Full details of any deviations from the Illumina 16S library preparation protocol must be supplied with data submissions.

3.15   The sequencer runs must backed-up onto local storage facilities, it is also recommended the runs are linked to Illumina's Cloud Storage (Basespace).

3.16   During and upon run completion, various run quality parameters should be monitored, including average Q30, which reflects average sequence quality and reliability as well as cluster density on the flow cell and the % of sequence reads passing quality filters.

3.17   Runs should be set up to automatically demultiplex the samples upon completion and fastq sequence files will be generated and stored in sample-specific folders for downloading from the Cloud for subsequent bioinformatics analyses.

3.18   For meta-barcode samples comprised of diverse bacterial communities, a sufficient quantity of high-quality, merged reads (ca 10000 reads) is required for the downstream derivation of IQI scores. Samples which fail to achieve enough quality reads may not receive an IQI score. In these cases, re-sequencing or a switch to invertebrate analysis will be required.

4. What to submit

4.1   Email aquaculture.monitoring@sepa.org.uk the following:
   1. A file sharing link to download the raw `.fastq.gz` DNA sequencing files.

11

2. Attach the corresponding 'Marine Pen Fish Farm Environmental Monitoring Survey Results' Excel spreadsheet

4.2   **NOTE** We recommend file sharing via Microsoft OneDrive or Sharepoint. Other online storage providers or servers can be used. However, if using a different system, please check with SEPA to verify cybersecurity standards before submission.

4.3   Ensure the 'eDNA Sample ID' fields within the attached Excel file matches the `YOUR-EDNA-SAMPLE-ID` applied to each pair of forward (`R1`) and reverse (`R2`) `.fastq.gz` files. These paired files, represent a single sample processed by the laboratory. The filenames should match the Illumina file format. For instance, `YOUR-EDNA-SAMPLE-ID_L001_R1_001` or `YOUR_EDNA-SAMPLE-ID_L001_R2_001`. Note the suffix (for example `_L001_R2_001`) is automatically added and should not be removed.

4.4   Once submitted, the bacterial DNA sequences will be compared against known bacteria codes held in international databases to enable the bacteria in the sample to be named and then used to derive an IQI score for the sample.

12

Annex 1: Seabed sediment types for which DNA can be used

Seabed sediment classifications are characterised based on the Folk Triangle (Folk, 1954). The classification of which is used as an acceptance criterion for using the model to derive an IQI score from DNA analyses.

Where sediment classifications for submitted data fall outside the range that the model was trained on, this acceptance criterion is not met. Instead, seabed invertebrate analysis must be undertaken.

The current acceptance range of sediment classifications is specified below

- gM - Gravelly Mud
- (g)M - Slightly gravelly mud
- (g)sM - Slightly Gravelly Sandy Mud
- sM - Sandy mud
- M - Mud

Where appropriate, SEPA will incorporate the additional data (matched DNA & invertebrate analyses) into the training dataset and periodically update the model. This will have the effect of extending the range of seabed sediment types in which DNA can be used to determine IQI scores.