FishGenome

"Improving cost-efficiency of fisheries research surveys and fish stocks assessments using next-generation genetic sequencing methods"

Technical guidelines to integrate genomic-based approaches into fisheries data collection

FishGenome - Deliverable 2.3

Rebeca Rodríguez, Laura Casas and Fran Saborido-Rey

[1] *Institute of Marine Research (IIM-CSIC), Vigo, Spain*

Corresponding author: rrodriguez@iim.csic.es

FishGenome is developed under the contract EASME/EMFF/2018/015: "Improving Cost-Efficiency of Fisheries Research Surveys and Fish Stocks Assessments using Next-Generation Genetic Sequencing Methods", financed by the Union's budget.

Table of contents

ACRONYMS

BITS	Baltic International Trawl Survey
CKMR	Close-kin Mark-Recapture
CTD	Conductivity, Temperature, and Depth device
DCF	Data Collection Framework
DNA	Deoxyribonucleic acid
DNAm	Epigenetic age determination
eDNA	Environmental DNA
EU	European Union
HTS	High-Throughput Sequencing
IBTS	International Bottom Trawl Survey
ICES	International Council for the Exploration of the Sea
MS	Member States
NS IBTS	North Sea International Bottom Trawl Survey
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
RAD-Seq	Restriction site-associated DNA sequencing
SNP	Single-nucleotide polymorphism
VMEs	Vulnerable Marine Ecosystems

CHAPTER 1: INTRODUCTION

The main purpose of this document is to provide guidance on how to integrate the genomic-based approaches tested in FishGenome into data collection carried out in bottom trawl surveys. Firstly, the data types and data collection process in bottom trawl surveys (as is usually carried out) is described. Then, a brief description of the genomic approaches that were tested in the pilot studies in FishGenome is presented. These approaches are: Close-kin mark-recapture (CKMR), Restriction site-associated DNA sequencing (RAD-Seq) for population connectivity and substructure, epigenetic age determination (DNAm) and environmental DNA (eDNA). The next section explains the main aspects to consider for the integration of genomic sampling into bottom trawl surveys, such as the requirements of the genomic method being considered and the survey characteristics (vessel capacity, available staff, etc.). Finally, the sampling experience from FishGenome is presented, including the sampling design and protocols followed for each genomic method. General recommendations are provided at the end of the document.

1.1 Overview of data collection in trawl-based research surveys

Member States (MS) in the EU regularly conduct research surveys of marine fishery resources to provide data for assessing the condition of exploited fish stocks and for monitoring the general status of the marine ecosystem. In the surveys, samples are taken throughout the potential range of the targeted fish species using standardized sampling gears, for example, with trawls and seines. At present, there is a list of mandatory surveys established in Commission Implementing Decision (EU) 2021/1168, and includes 51 surveys that are carried out in the Baltic Sea, the North Sea and eastern Arctic, the North Atlantic and the Mediterranean and Black Sea. From these, 19 are bottom trawl surveys carried out by EU MS and other European countries.

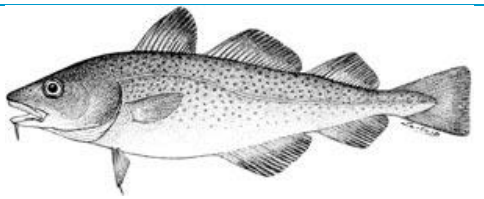
Survey data are important in stock assessment because they provide indices that help tuning the stock assessment models (e.g., the index of fish abundance, typically the number or weight of fish caught per unit of effort). For example, the Baltic International Trawl Survey (BITS) provides two indices (one from the BITS-Q1 and another from the BITS-Q4) used in the stock assessment of cod in the eastern Baltic Sea. Equally important, surveys provide biological information of fish populations, such as their size and age

distributions, the size-age relationships, estimates of the percentage of fish mature at each age, and information on their reproductive performance. Sampling programs in the surveys also include collection of data regarding the marine environment, such as the temperature and salinity of seawater. Actually, the sampling design used in surveys is usually based on a large number of species distributions and takes into account the existing communities and environmental aspects. It is important to note that the ability to collect data in bottom trawl surveys depends on several interrelated factors, such as the objectives of the survey, the fishing gear used, vessel size and storage capacity, the equipment available in the vessel and the available scientific and technical staff.

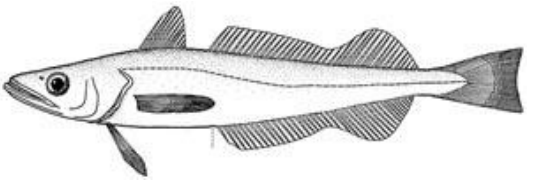
A summary of the types of data collected in in European bottom trawl research surveys at sea is presented in the Table 1. Common target species of bottom trawl surveys, for which biological sampling is carried out, are shown in Figure 1.

Type of data	Description
Species/biodiversity	Organisms are identified to the lowest practicable taxonomic level. Benthic organisms may also be identified (e.g., species, numbers and total weight per species per haul).
Abundance	Refers to the number of fish in a given fish population. Abundance estimations are based on the numbers of sampled fish for a species.
Length	Length distributions are recorded for all fish species caught and selected commercial cephalopods and shellfish
Total weight	Total weight for each species is recorded
Biological data	Usually for target species and other species of interest (e.g., elasmobranchs). Biological sampling usually includes individual weight, length, sex, maturity and age. Fish parasites and tissue samples for genetic analyses may be collected.
Diet	Stomach contents are collected to determine diet.
Hydrographic data	Temperature, salinity and oxygen are determined.
Marine litter	Marine litter is classified and recorded.
Other data	Fish eggs and plankton; fish tagging, sediment samples, water samples, interactions with marine mammals, seabirds and turtles; monitoring of non-indigenous invasive species; underwater images and video, acoustic for fish species, multibeam for seabed mapping

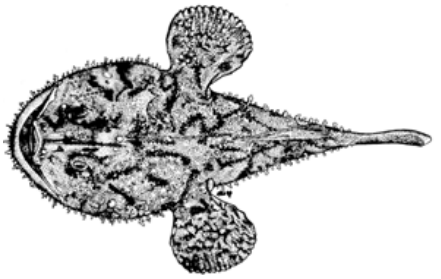
Table 1.- Summary of data collected in European bottom trawl research surveys at sea.



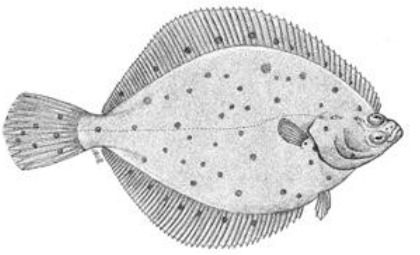
Cod (*Gadus morhua*)



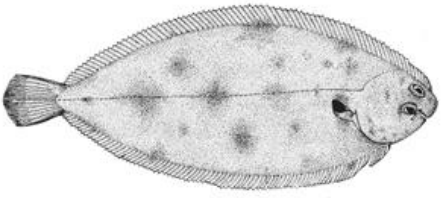
European hake (*Merluccius merluccius*)



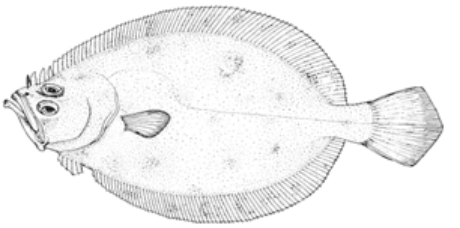
Monkfish (*Lophius* spp.)



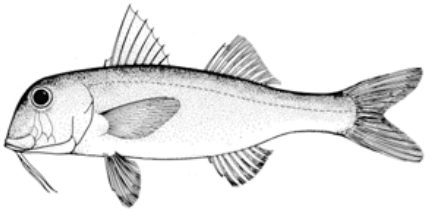
European plaice (*Pleuronectes platessa*)



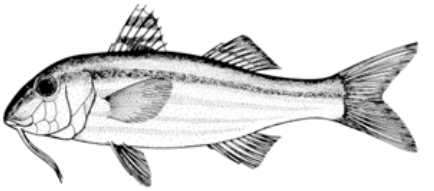
Sole (*Solea solea*)



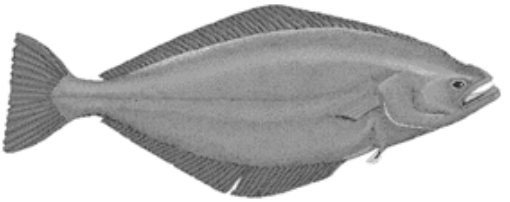
Megrim (*Lepidorhombus whiffiagonis*)



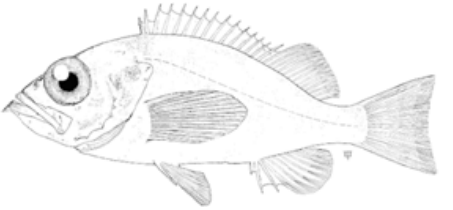
Red mullet (*Mullus barbatus*)



Striped red mullet (*Mullus surmuletus*)

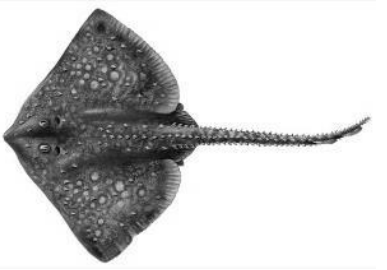


Greenland halibut (*Reinhardtius hipoglossoides*)

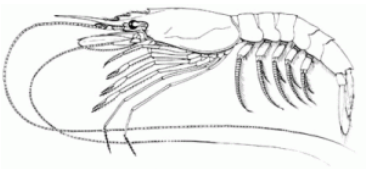


Redfish (*Sebastes* spp.)

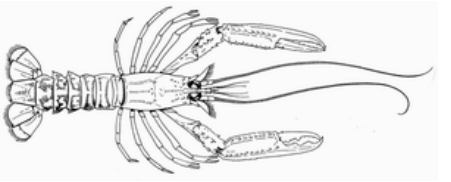
Sharks



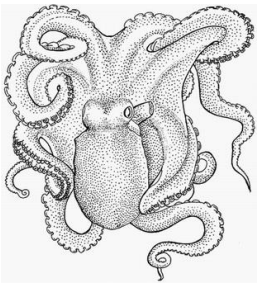
Rays



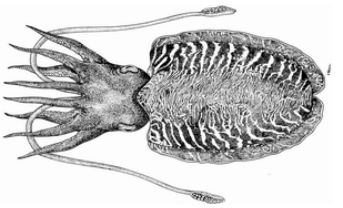
Shrimp (e.g., ***Aristeus antennatus***)



Norway lobster (***Nephrops norvegicus***)



Octopus (e.g., ***Octopus vulgaris***)



Cuttlefish (***Sepia officinalis***)

Squid (e.g., ***Loligo vulgaris***)

Figure 1.- Common target species for which biological information is routinely collected in bottom trawl surveys. Source: FAO.

1.2 Genomic methods

1.2.1 Close-Kin Mark-Recapture (CKMR)

Close-Kin Mark-Recapture (CKMR) is a new method for estimating abundance and other demographic parameters (e.g., population trend, survival rates, connectivity) from kinship relationships determined from genetic samples (Bravington, Grewe, and Davies 2016). CKMR offers a direct and fisheries-independent methodology to estimate abundance of wild fish stocks and therefore has the potential to be widely deployed for routine assessments of fisheries resources. The CKMR approach has been applied to a small number of fish species to date (e.g., brook trout) and, in cases where traditional independent estimations existed, a good agreement was found between abundance

estimates produced by this new methodology. In FishGenome, CKMR is being tested with three commercial stocks, hake (*Merluccius merluccius*) from the Mediterranean Sea and the North Sea and cod (*Gadus morhua*) from the North Sea.

1.2.2 Environmental DNA

In recent years, environmental DNA (eDNA) coupled with metabarcoding methodologies has emerged as a promising tool with the potential to improve biodiversity assessment, diet analysis, detection of rare or invasive species, population genetics, and ecosystem functional analysis (Bohmann et al. 2014; Goldberg et al. 2015). Taberlet et al. (2012a) define eDNA as a complex mixture of genomic DNA from many different organisms found in an environmental sample. This refers to any DNA that is collected from an environmental sample rather than directly from an organism, originating in cells from the body or waste products of organisms. This DNA is released from organisms into a variety of environmental samples such as soil, seawater, snow, or even air, and remains suspended in the water column or in the sediment (Ficetola et al. 2008). eDNA can be used to harness information encoded in seawater (Lodge et al. 2012; Taberlet et al. 2012a; Bohmann et al. 2014; Creer and Seymour 2017) and has been shown to be effective for identifying several organisms in marine and freshwater samples; it also shows promising potential for quantifying biodiversity in natural aquatic ecosystems. However, contrary to conventional methods, eDNA does not provide information on life stages, demographic structure of the population, fecundity, or health of the target species (Herder et al. 2014; Evans and Lamberti 2018). The FishGenome project is focused on two main methods to analyze eDNA: High-Throughput Sequencing (HTS) for biodiversity assessment, and quantitative Polymerase Chain Reaction (qPCR) for the quantification of a target species.

1.2.3 Epigenetics for age determination (DNAm)

DNA methylation changes occur as animals age and some of these changes are of a clock-like type. Consequently, identifying the exact loci that exhibit DNA methylation changes highly correlated with the pass of time is the basis for the development of epigenetic clocks. Their development in humans (Horvath, 2013) and other vertebrates (Paoli-Iseppi et al., 2019) suggested that epigenetic clocks should be possible to be developed in fish for age determination. The major doubt about the success of such

enterprise was the fact that all previous clocks in vertebrates were generated in several mammalian and a single bird species, i.e., warm-blooded species with determinate growth, while fish are cold-blooded and have indeterminate growth. Recently, the first epigenetic clock in fish was developed for the European sea bass, taking advantage of the fact that age could be known with accuracy because fish used were of hatchery origin (Anastasiadi and Piferrer, 2020). This achievement, therefore, paves the way for the development of similar clocks in other fish species of commercial importance. This constitutes, therefore, a very promising framework upon which additional applications across species can be developed. It is believed that developing piscine epigenetic clocks for target species could have a major impact since it will likely provide an accurate method for assessing age in fish and circumvent the limitations of the current methods (such as otolith readings). Importantly, improved age prediction would contribute to enhance fisheries management in a context of overexploited fish stocks worldwide. In FishGenome, an epigenetic clock is being developed for cod from the North Sea.

1.2.4 RAD-Seq for population connectivity and substructure

Restriction site-associated DNA sequencing (RAD-Seq) is a method that uses an Illumina sequencing platform to simultaneously discover and genotype tens to hundreds of thousands of single-nucleotide polymorphism (SNP) markers in hundreds of individuals for minimal investment of resources (Etter et al. 2011). These genetic markers can be used for evolutionary, phylogenomic and population structure studies among others (Díaz-Arce and Rodríguez-Ezpeleta, 2019 and references therein). In the fisheries context, genomics has been successfully used to define fish stocks and quantify the extent of adaptive divergence and connectivity between them, also allowing performing mixed-stock analysis with substantially increased resolution (Bernatchez et al., 2017). For example, differences of RAD-Seq-derived SNP frequencies within stocks could be analyzed to define their substructure (e.g., Leone et al 2019, Longo et al., 2021, Ceballos et al. 2021); or screening of SNPs could be used to search for sex markers that would allow sexing fish individuals (e.g., Palaiokostas et al. 2013; Gamble, 2016; Feron et al, 2021).

CHAPTER 2: DATA COLLECTION PROCESS IN TRAWL SURVEYS

As seen in the previous section, several types of data are collected in bottom trawl surveys. In bottom trawl surveys, data are collected for each haul that is carried out. The haul starts at the moment when the gear starts fishing on the seabed and ends when the gear starts being pulled back. For example, in the North East Atlantic bottom trawl surveys, haul duration varies from 20 minutes to 60 minutes (ICES, 2017). The first data collected are the **haul characteristics** such as location, depth, direction and speed, warp length, vertical and horizontal opening of the net, door spread, net symmetry and behavior, etc. Once the haul is finished and the net is retrieved, the catch is processed. Data collection of the catch is carried out by the scientific and technical staff onboard.

2.1 Catch sorting and sampling

In general, the process for sampling trawl catches consists of the following steps:

1. **Sorting the catch.** Generally, the entire is sorted. All fish, crustaceans and molluscs species are identified to the lowest practicable taxonomic level that is possible in the field. When it is not possible to sort the entire catch, for example, when very large catches occur, adequate subsamples are taken before proceeding to the taxonomic identification of organisms. In many national surveys, other benthic organisms such as sponges, corals and sea pens caught in the gear are sorted as well. Although trawl gears used in the different surveys are not effective for catching benthos for quantitative sampling, these catches help obtaining information about the composition and distribution of Vulnerable Marine Ecosystems (VMEs)[1]. If a live specimen of a rare species or a species subject to conservation measures (e.g., many sharks and rays) is caught, the animal should be released as quickly as possible after obtaining its length, weight and sex.

[1] Vulnerable marine ecosystems (VMEs) are groups of species, communities, or habitats that are vulnerable to adverse impacts from fishing activities. They possess traits such as slow growth rates, long life expectancies and low or unpredictable recruitment. Examples of organisms that constitute VMEs are cold water corals, deep sea sponge communities and seep and vent communities.

2. **Recording total weight** for each species.

3. **Measuring and recording length.** Length distributions are recorded for all fish species caught and selected commercial cephalopods and shellfish. Length is usually defined as total length, which is measured from the tip of snout to tip of the end of the caudal fin. For most fish species (and also for cuttlefish and squid species), length is measured to 1 cm below. For smaller fish such as sardine and anchovy, length is measured to the 0.5 cm. Crustaceans are measured to 0.1 cm below. In the case of elasmobranch species, they are usually measured and weighed by sex.

4. **Obtaining biological data if required**. This is done only for target species or specific species determined in the survey protocol. Given the large area covered by the surveys (e.g., Northeast Atlantic, Mediterranean, Northwest Atlantic, North Sea and Baltic Sea) and the number of countries participating in such surveys, the list of target species varies between surveys and countries. If the species is required for further biological data, then weights (e.g., total weight and gutted weight), sex and maturity are recorded. Ageing material (e.g., otoliths and illicia[2]) is obtained and stored for age determination at the lab. In many surveys, stomach contents are also collected to determine diet. Other biological samples can be obtained in this step, for example, tissue samples for genetic analyses and parasites samples can be collected.

[2] In anglerfish, the first ray of the first dorsal fin is known as illicium. It is used by anglerfish as bait to lure their prey. It is a thin, flexible and very mobile structure, larger than the following rays, ending in a skin lobe, (Bauchot and Pras, 1987). It is considered as a better skeletal structure for age estimation, because better results were obtained than using otoliths.

Figure 2.- Sorting the catch, measuring length and biological sampling in a survey.

2.2 Environmental data

Normally, immediately before or after each fishing haul, hydrographical data are collected. Such data include sea surface and bottom temperature and salinity. Usually, a Conductivity, Temperature and Depth device (CTD) is used for measuring temperature and salinity. CTDs are frequently attached to a large metal frame known as rosette or carousel, which may hold several bottles that are used to collect water at different depths, as well as other sensors that can measure additional physical or chemical properties, for example, chlorophyll, turbidity (the amount of particles in the water), irradiance (amount of light from the sun that penetrates the water), and the amount of dissolved oxygen in the water (Figure 3). Other additional data that may be recorded are surface current speed, bottom current direction, bottom current speed, wind direction, wind speed, and swell height and direction.

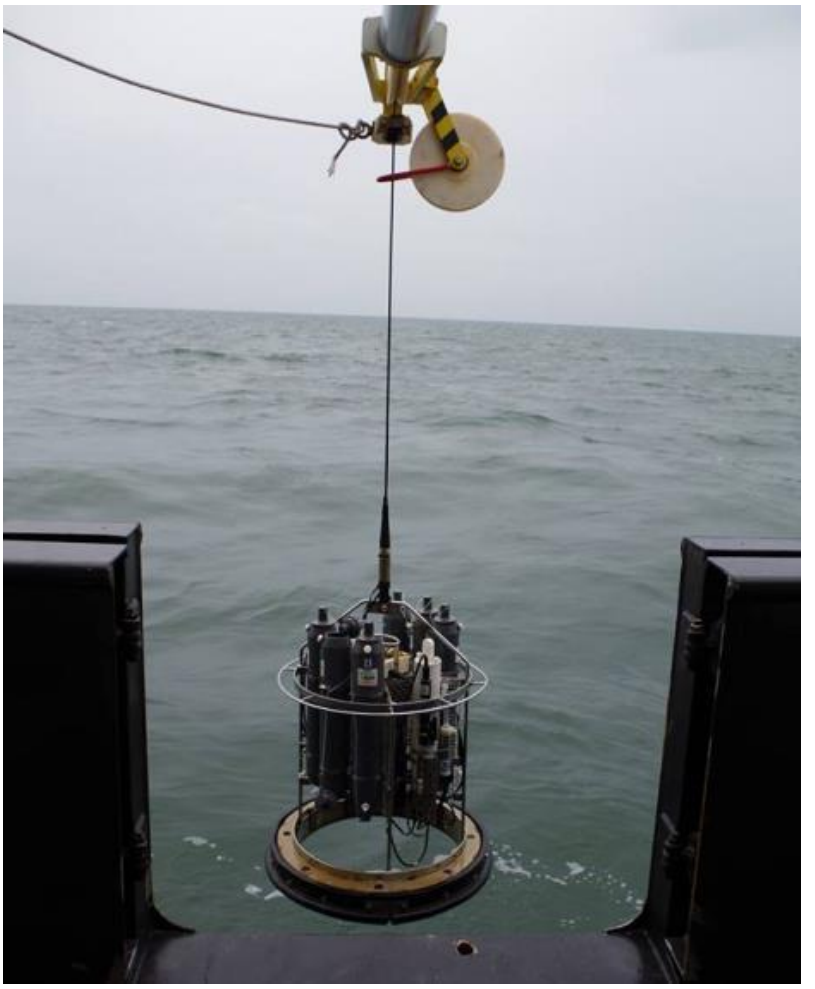


Figure 3.- Image of CTD integrated in a rosette containing water-sampling bottles (Niskin bottles). Source: Flanders Marine Institute (VLIZ).

2.3 Sediment samples

Sediment is particulate matter that can be transported by physical processes (e.g., wind or water currents) and is gradually deposited on the seabed. It originates from a variety of sources, for example, organic material coming from marine organisms, eroded material from land transported to the ocean by rivers, ice or wind, volcanic ash, etc. Sediment samples are commonly taken in research surveys, because these samples allow studying many aspects of the seabed, such as its type and composition, the presence of contaminants, climatic past events and the organisms that live on or in the seabed (benthos).

In bottom trawl surveys, the seafloor characteristics (type of sediment) are usually recorded and sampling of benthic communities is commonly carried out. Sediment samples may be collected using a wide variety of methods and equipment, depending on several factors, such as water depth, the type of sample required, the volume of sediment required, the depth of the sediment penetration required and the sediment characteristics (grainsize, competence, porosity, etc.). There are several types of sediment sampling devices, but grab samplers (e.g., Ekman and Van Veen grabs), and core samplers (e.g., gravity corers, piston corers and box corers) are most commonly used in research surveys. Dredge samplers are also used to collect sediment and benthic organism. However, since these devices scrape along the surface of the sediment there is considerable disruption to the sediment. In some occasions, Remotely Operated Vehicles (ROVs) can be also used to collect sediment samples.

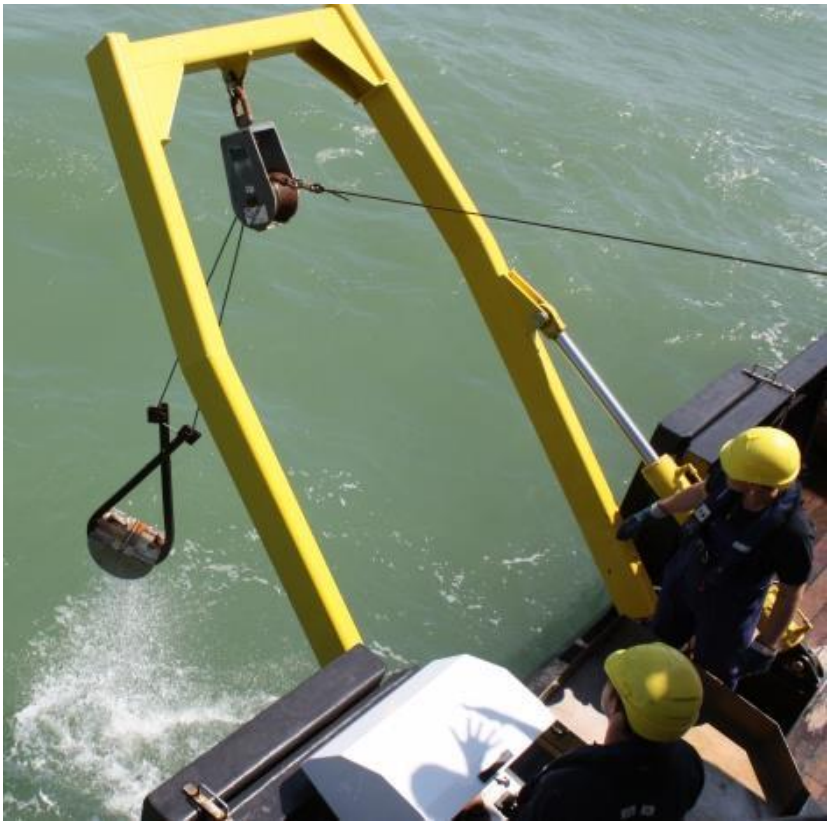


Figure 4.- Image of a Van Veen grab used to collect sediment samples. Source: Flanders Marine Institute (VLIZ) - Mertens, Tina).

CHAPTER 3: INTEGRATION OF GENOMIC SAMPLING INTO DATA COLLECTION IN BOTTOM TRAWL SURVEYS

Several aspects need to be considered to integrate genomic sampling in bottom trawl surveys. Firstly, we need to have clear goals regarding the genomic study that we want to carry out, and the following questions can help clarify this:

- What are the goals of the study?
- What type of genomic technique is the most adequate for the study?
- What type of samples are needed?
- How many samples are needed for the study?

Then, the survey characteristics need to be considered. The following question can be used to determine if the survey being considered is adequate:

- Does the survey cover the right species and areas?
- Does the vessel characteristics allow the sampling to be carried out? Consider here the vessel capacity, storage facilities in the vessel (such as availability of freezers), vessel equipment (e.g., grabs/ corers for sediment samples, CTD carousel water sampler with Niskin bottles), etc.
- Does the survey protocol allow for extra work? Is it possible to integrate the sampling into the existing survey protocols?
- Is there enough scientific and technical staff to carry out the sampling? If not, is it possible for dedicated staff to go on board to carry out genomic sampling?

Once these aspects are figured out, specific sampling protocols for each genomic technique must be developed. These protocols must be tailored to the survey in which they are going to be implemented, and should be integrated as much as possible into the existing workflow and sampling protocols of the survey. The protocols must consider the material and equipment needed, the methodology to obtain samples and the conservation and storage conditions of the samples (after collection, during transportation, etc.).

One should bear in mind that tissue samples for genomic analyses are often collected during research surveys (although not in a routine manner) and hence, no difficulties would be expected for this. Normally, genetic samples can be collected during the biological sampling of specimens (step 4 of "Catch sorting and sampling" in section 2).

However, collecting samples for eDNA analyses (water and sediment) requires devices and protocols that are not always available in all research vessels. Another consideration regarding eDNA samples is that they should be collected immediately before the haul, or some hours in advance, in the same location where the haul is planned (e.g., the afternoon before, if the haul is taking place the next morning). This is because during the haul, fish caught in the net releases DNA to the water and also the sediments are re-suspended during the fishing operation. This can produce biases in the quantification of DNA using eDNA. Also, one must take into account that genomic techniques are highly sensitive to potential contamination of samples during collection and manipulation. Therefore, special care must be taken during the samplings to avoid contamination and reduce the rate of erroneous assignments.

CHAPTER 4: SAMPLING EXPERIENCE IN FISHGENOME

To illustrate the process of integrating genomic sampling into bottom trawl surveys, the sampling experience in FishGenome is described here. The first subsection explains the experimental design followed, the second subsection contains the sample collection protocols for each genomic technique (i.e., CKMR and RAD-Seq, epigenetic age determination and eDNA), and the third subsection offers some general comments about the sampling process in the surveys.

FishGenome samplings were carried out in two pilot studies in two bottom trawl surveys during 2019, namely, the **MEDITS survey** (GSA5; Balearic Islands) and the German part of the **North Sea IBTS Q3**. The GSA5 MEDITS survey is carried on R/V "Miguel Oliver" (Overall length 70 m, gross tonnage: 2495, capacity for 22 crew members and 23 scientists/technicians, six labs onboard). The NS IBTS Q3 survey (German part) is carried on R/V "Walther Herwig III" (Overall length 63.18 m, gross tonnage: 2131, Capacity for 21 crew members and 12 scientists/ technicians, seven labs onboard) (Figure 5).

The sample requirements for FishGenome consisted of:

1. **Tissue samples** for cod and hake in the North Sea survey and hake in the MEDITS survey for the CKMR for abundance estimation and RAD-Seq analyses for population connectivity and substructure.
2. **Water and sediment samples** to conduct eDNA analyses.

A single agreed protocol was established for the two surveys with clear indication of experimental design, sampling intensity, sample collection, preservation, data gathering and other details that ensure traceability and usability of the samples for the FishGenome goals.

A

B

Figure 5.- Research vessels used in the MEDITS GSA5 survey: R/V MiguelOliver (A) and NS IBTS Q3 (German part): R/V Walther Hewig III (B) Source: Ministerio de Agricultura, Pesca y Alimentación and Thuenen Institut, respectively.

4.1 Sampling designs

The sampling designs for each genomic technique were as follows:

For **CKMR/ RAD-Seq**, the goal of the sampling was to collect a minimum total number of 600 specimens in the North Sea, which corresponds to sampling virtually all the fish captured, considering the 2018 survey catches (490 specimens of North Sea cod and 188 specimens of North Sea hake). Grouping the individuals by 5-cm size classes, up to 150 individuals by size range needed to be collected. In the Mediterranean, the goal was to collect in total 750 individuals, between 150-200 individuals by size range, grouping the sizes by 5-cm classes (the MEDITS-GS5 2018 survey accounted for 1042 hake specimens in 2018). The following tables summarize the number of samples for each species/area and by size range.

NORTH SEA COD

North Sea cod – number of individuals	
Size range	**Fin clip**
<25	150*
25-29	150*
30-34	150*
35-39	150*
40-44	150*
45-49	150*
>50	150*

*Since the goal is to collect 600 individuals in total, more than 150 fish can be collected in the most abundant size-class, although it is highly unlikely.

NORTH SEA HAKE

North Sea hake – number of individuals	
Size range	**Fin clip**
<30	150*
30-39	150*
40-49	150*
50-59	150*
60-69	150*
70-79	150*
>80	150*

*Since the goal is to collect 600 individuals in total, more than 150 fish can be collected in the most abundant size-class, although it is highly unlikely.

MEDITERRANEAN HAKE

Mediterranean hake – number of individuals	
Size range	**Fin clip**
<10	150
10-14	150
15-19	200
20-24	200
25-29	All**
30-34	All**
>35	All**

**The probability of catching 150 individuals larger than 25 cm is very low (in 2018, only 80 individuals were captured), that is the reason for the intensive sampling in lower size ranges.

The sampling design for **epigenetic age determination** requires an even distribution of age groups and both sexes evenly represented. In total, five time-points represented by 10 individuals (five males + five females) were required for the analysis. Since during the sampling age of the fish is still unknown (otoliths are collected but not processed) the size was used as a proxy for the time-points. The goal of the sampling was to collect the fin clip, liver, branchial arch and muscle of 20 fish by size range and sex. Although some of the smaller undifferentiated specimens did not allow sexing during the sampling, they were collected anyway.

NORTH SEA COD

Cod - number of individuals						
	Liver		**Branchial arch**		**Muscle**	
Size range	♂	♀	♂	♀	♂	♀
<25	20		20		20	
25-29	10	10	10	10	10	10
30-34	10	10	10	10	10	10
35-39	10	10	10	10	10	10
40-44	10	10	10	10	10	10
45-49	10	10	10	10	10	10
>50	10	10	10	10	10	10

MEDITERRANEAN HAKE

Mediterranean hake - number of individuals						
	Liver		Branchial arch		Muscle	
Size range	♂	♀	♂	♀	♂	♀
<10	20		20		20	
10-14	20		20		20	
15-19	10	10	10	10	10	10
20-24	10	10	10	10	10	10
25-29	10	10	10	10	10	10
30-34	10	10	10	10	10	10
>35	All	All	All	All	All	All

Regarding **eDNA water and sediment samples**, the aim was to collect water and sediment samples, using a carousel water sampler and a grab, respectively. Water samples were collected at trawl depth before trawling: either right before the trawl (in the NS IBTS survey) or the afternoon before (in the MEDITS survey). In the MEDITS survey, some flexibility was considered, and samples could be also collected right before the trawl, in order to have information to compare both approaches (eDNA *vs* trawling).

In the MEDITS survey, the available sampling equipment was a carousel water sampler with 5L Niskin bottles (capacity of 12 Niskin bottles) and a CTD. At each sampling station, two Niskin bottles (2x5L) needed be collected to have a replica for each sample. 14 sampling stations in total were targeted, concentrating on the areas where higher biomasses of hake/cod are usually found. Bottles needed to be rinsed with distilled water in between samples to avoid cross-contamination. Two people from the scientific/technical staff were assigned to carry out these samplings. In the NS IBTS survey, the available equipment was a carousel water sampler with 1.7L Niskin Bottles and a CTD. Volume from several bottles were pooled to obtain 5L (+ 5L replica). The aim for sediment samples was to collect them at the same stations used for water collection whenever possible. In both surveys, sediment sampling was carried out using a Van Veen grab.

4.2 Specific guidelines/protocols for each technique

The following sampling protocols were designed tailoring the specific needs of the research surveys where FishGenome samplings were carried out, namely, the MEDITS (GSA5; Balearic Islands) and the German part of the North Sea IBTS Q3. One of the most important considerations when designing these protocols was to allow their implementation without disturbing the standard procedures carried out on board during the regular research surveys. For these reason, application to other surveys may require small adjustments of the sampling protocols.

4.2.1 CKMR and population connectivity and substructure using RAD-Seq

An ideal Close Kin Mark Recapture (CKMR) sampling design for estimating population size should have an even distribution of both reproductively mature and juvenile groups and both sexes evenly represented. For connectivity and fine-scale substructure using RAD-Seq, two or more different stocks/populations/groups need to be sampled while the search of a sex marker requires the presence of individuals of both sexes.

SAMPLING COLLECTION

Genetic information for CKMR, connectivity, fine scale-substructure and isolation of a sex marker is obtained through the extraction of DNA from any tissue sampled from individual fish. However, fin clips are generally used due to their non-invasive nature. Large numbers of individuals are routinely sampled during the course of traditional research trawl surveys and our proposal adds a simple, inexpensive, and time efficient addition to most sampling protocols already in place.

EQUIPMENT AND SUPPLIES

1) Fish Measuring Board
2) Electronic balance and tub
3) 1.5-2ml Eppendorf or screw-cap tubes
4) Tube rack holding tubes
5) Tissue Scissors or sterile, disposable scalpels
6) Forceps
7) Clipboard
8) Data Entry Sheet
9) Pencil
10) Squeeze Bottle filled with molecular grade ethanol ≥ 95%

METHOD

A detailed step-by-step protocol is detailed below and depicted in Figure 6.

1. Record date, location and morphological data (i.e., total length, weight, etc...) together with all other relevant biological information if available (e.g. sex, reproductive condition, etc…). Collect otoliths for age determination. Label the tube with a pencil.

2. Using a scissor or a disposable scalpel, remove a small piece of fin (2 cm^2/2 cm x 1 cm or about the size of your thumb).

3. Use the forceps to transfer the fin clip to a tube and label the vial with a pencil either inside or outside. Fill the tube with ≥ 95% ethanol, making sure that the ratio between ethanol and tissue is at least 3:1 and the tissue is completely covered by ethanol.

4. Rinse and wipe the instruments (scissors and forceps) after every use to ensure that remnants of tissue are not present before dissecting the next specimen to avoid cross-contamination (or keep them in a glass filled with ethanol in between samples).

5. Tissue samples preserved in ≥ 95% ethanol can be stored at room temperature, refrigerated, or frozen at -20°C. Avoid direct exposure to sunlight.

6.- For long-term, store at -20°C.

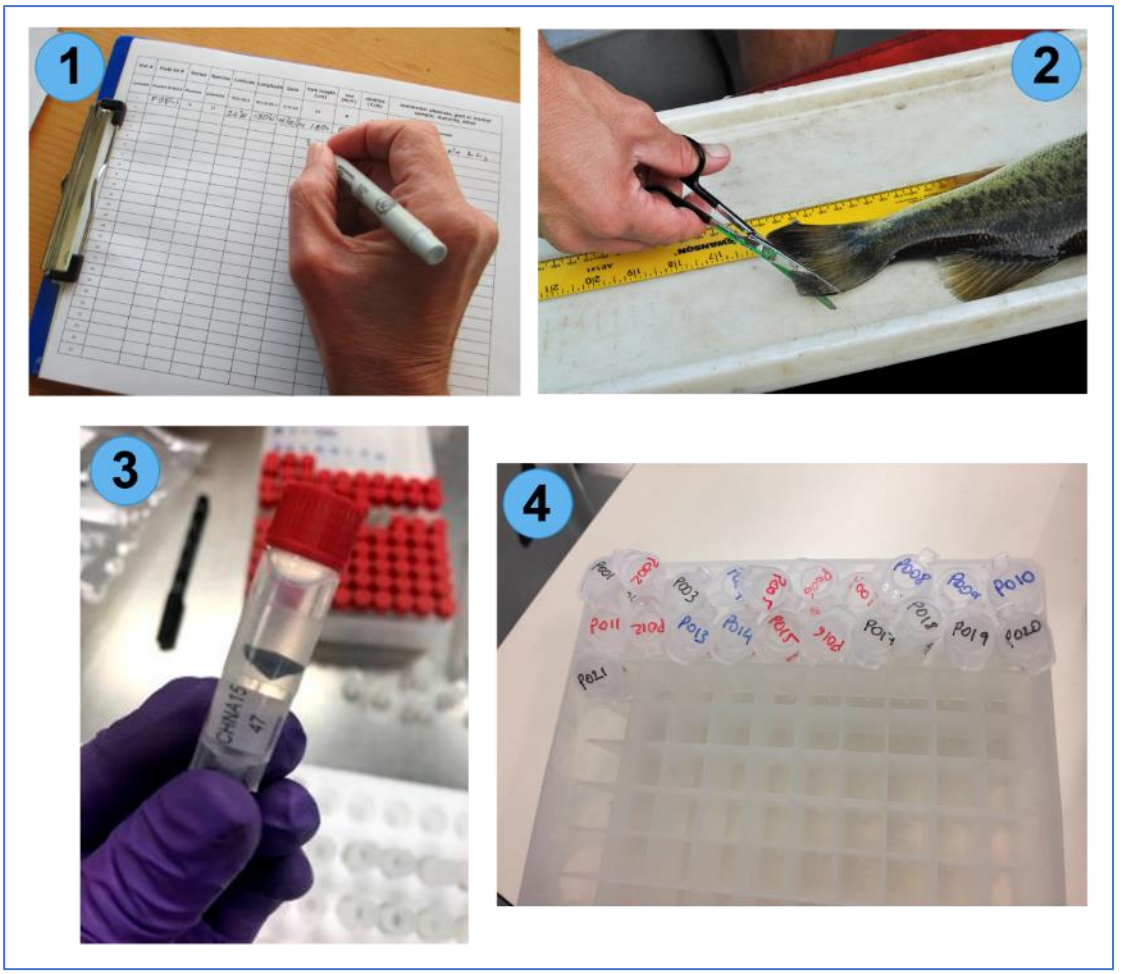


Figure 6.- Schematic sampling for CKMR. 1. Record total length, date, location and all other relevant biological information if available (e.g. sex, reproductive condition, if otoliths were obtained). 2. Removal of small piece of fin, 3. Place in ethanol, 4. Place in numbered vial.

4.2.2 Epigenetic age determination (DNAm)

SAMPLING DESIGN

An ideal DNAm sampling design should have an even distribution of age groups (size will be used) and both sexes evenly represented.

SAMPLING COLLECTION

Genetic information for DNAm age is obtained through the extraction of DNA from any tissue sampled from individual fish. At least two tissues should be collected for comparison, in our case, four tissues will be collected: fin clip, a piece of muscle, branchial arch and liver. Tissues should be dissected perimortem.

Equipment and supplies:

1) Fish Measuring Board
2) Electronic balance and tub
3) 1.5-2ml Eppendorf or screw-cap tubes
4) Tube rack holding tubes
5) Tissue Scissors or sterile, disposable scalpels
6) Forceps
7) Clipboard
8) Data Entry Sheet
9) Pencil/Write
10) Squeeze Bottle filled with NON-DENATURED/NON-METHYLATED molecular grade ≥ 95% ethanol

Method:

A detailed step-by-step protocol is detailed below and depicted in Figure 7.

1. Record date, location and morphological data (i.e., total length, weight, etc...) together with all other relevant biological information if available (e.g. sex, reproductive condition, etc…). Collect otoliths for age determination. Label the tube with a pencil.

2. Using a scissor or a disposable scalpel, remove a small piece of fin (2 cm^2/2 cm x 1 cm or about the size of your thumb).

3. Use the forceps to transfer the fin clip to a tube and label the vial with a pencil either inside or outside. Fill the tube with ≥ 95% NON-METHYLATED ethanol, making sure that the ratio between ethanol and tissue is at least 3:1 and the tissue is completely covered by ethanol.

4. Wipe the instruments (scissors and forceps) to ensure that remnants of tissue are not present before dissecting the next tissue to avoid cross-contamination.

5. Take a slice of muscle tissue (without skin) (2 cm^2/2 cm x 1 cm or about the size of your thumb) on the left side of the fish, below the dorsal fin.

6. Use the forceps to place the muscle in a tube and label the vial with a pencil.

7. Wipe the instruments (scissors and forceps) to ensure that remnants of tissue are not present before dissecting the next tissue to avoid cross-contamination

8. Remove a piece of branchial arch using a scissor or a disposable scalpel.

9. Use the forceps to place the branchial arch in a tube and label the vial with a pencil.

10. Wipe the instruments

11. Remove a small piece of the liver using a scissor or a disposable scalpel.

12. Use the forceps to place the liver in a tube and label the vial with a pencil.

13. Rinse and wipe the instruments (scissors and forceps) after every use to ensure that remnants of tissue are not present before dissecting the next specimen to avoid cross-contamination (or keep them in a glass filled with ethanol in between samples).

14. Tissue samples preserved in ≥ 95% NON-METHYLATED ethanol can be stored at room temperature, refrigerated, or frozen at -20°C. Avoid direct exposure to sunlight.

15.- For long-term, store at -20°C.

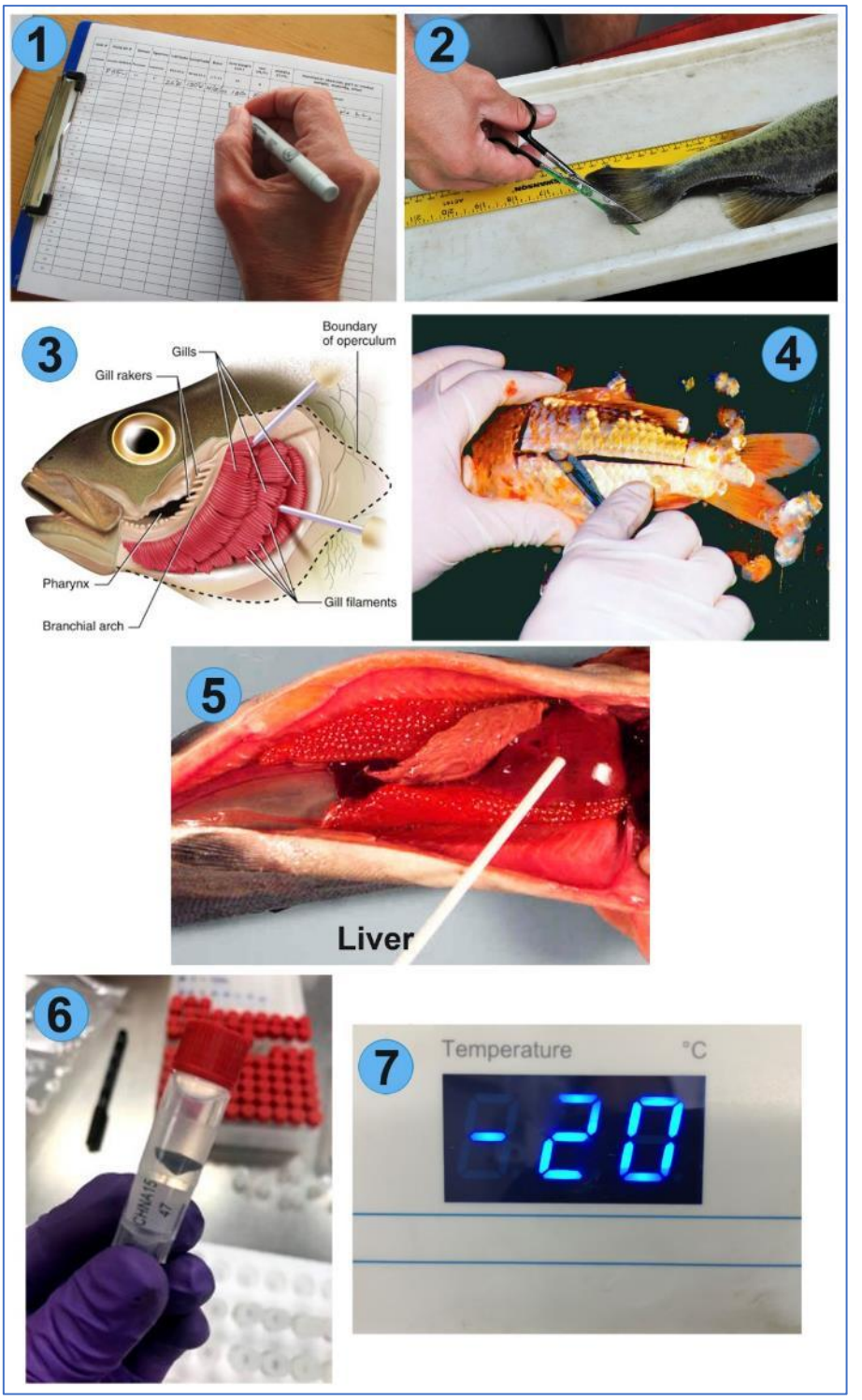


Figure 7.- Schematic representation of biological tissue sampling. 1. Record total length, date, location and all other relevant biological information if available (e.g. sex, reproductive condition, if otoliths were obtained); 2. Removal of small piece of fin; 3. Removal of muscle tissue; 4 Removal of piece of branchial arch; 5. Removal of tip of the liver; 6. Place in ethanol, 7. Store at -20oC for long-term.

4.2.3 eDNA

The eDNA method consists on the analysis of environmental samples to capture DNA from aquatic organisms. Two types of samples are going to be collected:

1.- Water samples: Should be collected at trawl depth before trawling, the afternoon or morning before the trawl to avoid disturbing the standard procedures carried out on board during the regular Fisheries research survey. Two replicas (2 Niskin 5L botlles) should be processed per sampling site.

2.- Sediment samples: Should be collected after the water. The sediment should be collected from the superficial strata in duplicates (2x per station) using a common corer or grab (e.g., Van Veen grab). The sampling device must be rinsed with sterile water in between samples.

SAMPLING COLLECTION

Equipment and supplies:

1) Carousel Water Sampler with Niskin Bottles and CTD
2) Filtration ramp equipped with a peristaltic pump and a compressor
3) Silicon tubing
4) Membranes with a pore size of 0.45 µm (MF-Millipore Membrane, MERCK)
5) Common corer
6) Sterile 2 ml tubes with silica gel tubes for membranes
7) Tweezers
8) 100 ml Sterile plastic containers for sediment
9) Clipboard
10) Data Entry Sheet
11) Pencil/Permanent marker
12) Molecular grade ethanol ≥ 95%
13) Bleach

Method:

WATER SAMPLES

This protocol involves filtering the water to capture the DNA from aquatic organisms and should be performed on board as soon as possible after water collection. If filtration cannot be performed *in situ*, water samples should be stored at -20ºC.

A detailed step-by-step protocol is detailed below:

1. Assemble filtration ramp and filter

2. Filter 5L of water per sampling station (2x for replicas) through a membrane with a pore size of 0.45 µm.

3. Once the filtration process of the 5L is finished, use sterile forceps to place the membrane in a sterile 2 ml thread tube containing silica gel.

4. Disinfect forceps and silicon tubing by soaking in 50 percent bleach solution for at least 1 minute and rinsing thoroughly in distilled water.

5. Store samples at -20ºC.

SEDIMENT SAMPLES

A detailed step-by-step protocol is detailed below:

1. Take approximately 50 gr (≈ 40 ml) of the superficial strata of the sediment with a plastic spoon (1 per sample, do not reuse) and place it in a 100 ml sterile container

2. Label with a pencil.

3. Add 60 ml of ≥ 95% ethanol.

4. Store at -20 °C.

4.3 Comments regarding the integration of genomic samplings in the MEDITS and NS IBTS surveys

In the MEDITS GSA5 survey the samplings were carried out according to plan with no major incidences or disruptions to the survey workflow. According to the survey report (Thuenen Institute, 2019) during the 2019 NS IBTS Q3 survey (German part), the sampling of water and sediments could be conducted without additional ship time, because DNA samples could be obtained from routine samples. Fish sampling was conducted partly on regular sampling stations (without additional ship time) and partly on five dedicated stations for genetic analyses. The latter amounted to a total of ca. 5 hours' additional ship time. Processing of water samples aboard required one additional scientist, and support from a second person for the days of the dedicated trial stations for genetics. Additional scientist and technician time required for the processing of fish samples was significant and turned out to involve two persons. Processing times per fish were between 5 and 10 minutes, depending on the number of different tissues sampled.

According to the experience in the FishGenome project, integrating the tissue sample collection for genomic analyses into the regular scientific fisheries surveys seems to be feasible in surveys with sufficient resources (i.e., staff and equipment), such as MEDITS and NS IBTS.

CHAPTER 5: GENERAL RECOMMENDATIONS

- A careful planning and sampling design, taking into account both the requirements of the genomic methods and the survey characteristics should be carried out to successfully integrate genomic sampling into data collection in bottom trawl surveys.

- In the sampling design, the number of samples must be balanced between the genomic methods requirements and the extra effort that this sampling implies. The number of samples must be adequate to not exceed the capacity of collection during the limited survey time.

- Specific sampling protocols for each genomic technique must be developed. These protocols must be tailored to the survey in which they are going to be implemented, and should be integrated as much as possible into the existing workflow and sampling protocols of the survey.

- Sampling protocols should include the standard guidelines regarding the data to be collected for each individual sample, with the purpose of obtaining a high-quality material suitable for genomic analyses.

- The area where biological samplings are carried out in research surveys are considerably limited in terms of space and it is not easy to always maintain cleanliness. Genomic techniques are highly sensitive to potential contamination of samples during collection and manipulation. Therefore, special care must be taken during such samplings to avoid contamination.

- Water and sediment samples must be collected before the haul is carried out, because during the haul, fish caught in the net releases DNA to the water and also the sediments are re-suspended during the fishing operation. This can produce biases in the quantification of DNA.

REFERENCES

Anastasiadi, D. & Piferrer, F. (2020). A clockwork fish: Age prediction using DNA methylation-based biomarkers in the European seabass. Molecular Ecology Resources. 20: 287–397. doi: 10.1111/1755-0998.13111.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M., Beacham, T. D., et al. (2017). Harnessing the power of genomics to secure the future of seafood. Trends in Ecology and Evolution, 32(9), 665-680.

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, et al. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution, 29(6): 358-367. DOI: 10.1016/j.tree.2014.04.003

Bravington, M. V., Grewe, P. M. & Davies, C. R. (2016). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. Nature Communications 7, 13162.

Ceballos, S. G., Papetti, C., Babbucci, M., Fernández, D. A., Schiavon, L., and Cheng, C. H. C. (2021). Genome-wide analysis reveals striking lack of genetic differentiation over long distances for the Antarctic toothfish Dissostichus mawsoni: High genetic connectivity or shared spawning grounds? Fisheries Research, 243, 106074.

Creer S and Seymour M. (2017) Marine ecology: genetics from a drop in the ocean. Nature Ecology & Evolution, 1(1): 0037. DOI: 10.1038/s41559-016-0037.

Díaz-Arce, N. and Rodríguez-Ezpeleta, N., (2019). Selecting RAD-Seq data analysis parameters for population genetics: the more the better? Frontiers in genetics, 10, p.533.

Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A. and Cresko, W.A., (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. Molecular methods for evolutionary genetics, pp.157-178.

Evans NT and Lamberti GA. (2018) Freshwater fisheries assessment using environmental DNA: A primer on the method, its potential, and shortcomings as a conservation tool. Fisheries Research, 197: 60-66. DOI: 10.1016/j.fishres.2017.09.013

Feron, R., Pan, Q., Wen, M., Imarazene, B., Jouanno, E., Anderson, J., et al. (2021). RADSex: a computational workflow to study sex determination using Restriction Site-Associated DNA Sequencing data. Molecular Ecology Resources, 21(5), 1715-1731.

Ficetola GF, Miaud C, Pompanon F and Taberlet P. (2008) Species detection using environmental DNA from water samples. Biology letters, 4(4): 423-425. DOI: 10.1098/rsbl.2008.0118

Gamble, T. (2016). Using RAD-seq to recognize sex-specific markers and sex chromosome systems. Molecular Ecology 25, 2114–2116. doi:10.1111/mec.13648

Goldberg CS, Strickler KM and Pilliod DS (2015) Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. Biological Conservation, 183: 1-3. DOI: 10.1016/j.biocon.2014.11.040

Herder J, Valentini A, Bellemain E, Dejean T, Van Delft JJCW, et al. (2014). Environmental DNA. A review of the possible applications for the detection of (invasive) species. Netherlands Food and Consumer Product Safety Authority. Stichting RAVON, Nijmegen. Report 2013-104. DOI: 10.13140/RG.2.1.4002.1208

Horvath, S. (2013). DNA methylation age of human tissues and cell types. Genome Biology, 14(10), R115. Retrieved from http://genomebiology.com/2013/14/10/R115/

ICES (2015). Manual for the International Bottom Trawl Surveys. Series of ICES Survey Protocols SISP 10 - IBTS IX. 86 pp.

ICES (2017). Manual of the IBTS North Eastern Atlantic Surveys. Series of ICES Survey Protocols SISP 15. 92 pp. http://doi.org/10.17895/ices.pub.3519

Leone, A., Álvarez, P., García, D., Saborido-Rey, F. and Rodriguez-Ezpeleta, N. (2019). Genome-wide SNP based population structure in European hake reveals the need for harmonizing biological and management units. ICES Journal of Marine Science, 76(7), pp.2260-2266.

Lodge DM, Turner CR, Jerde CL, Barnes MA, Chadderton L, et al. (2012). Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. Molecular Ecology, 21(11): 2555-2558. DOI: 10.1111/j.1365-294X.2012.05600.x

Longo, G.C., Harms, J., Hyde, J.R., Craig, M.T., Ramón-Laca, A. and Nichols, K.M. (2022). Genome-wide markers reveal differentiation between and within the cryptic sister species, sunset and vermilion rockfish. Conservation Genetics, 23(1), pp.75-89.

Palaiokostas, C., Bekaert, M., Davie, A., Cowan, M.E., Oral, M., Taggart, J.B., Gharbi, K., McAndrew, B.J., Penman, D.J. and Migaud, H., 2013. Mapping the sex determination locus in the Atlantic halibut (Hippoglossus hippoglossus) using RAD sequencing. BMC genomics, 14(1), pp.1-12.

Paoli-Iseppi, R. D., Deagle, B. E., Polanowski, A. M., McMahon, C. R., Dickinson, J. L., Hindell, M. A., & Jarman, S. N. (2019). Age estimation in a long-lived seabird (Ardenna tenuirostris) using DNA methylation-based biomarkers. Molecular Ecology Resources, 19(2), 411–425. doi: 10.1111/1755-0998.12981

Taberlet P, Coissac E, Hajibabaei M and Rieseberg LH (2012a) Environmental DNA. Molecular Ecology, 21(8): 1789-1793. DOI: 10.1111/j.1365-294X.2012.05542.x

Thuenen Institut (2019). Survey Report FRV Walther Herwig III, WH 428 08 July to 03 August 2019 - IBTS and GSBTS. Federal Research Institute for RuralAreas, Forestry and FIsheries. Thünen Institute of Sea Fisheries. 9 pp.