Ref. Ares (2020)1980204 – 08/04/2020

AtlantECO

Atlantic Ecosystems Assessment, Forecasting & Sustainability

Deliverable number 4.2 - Handbook of Standards and Best Practices Version 2

Dissemination level: Public

Related Work Package	WP4 Augmented observations and technological innovations
Related task(s)	Task 4.1.1
Lead beneficiary	European Molecular Biology Laboratory (EMBL-EBI)
Author(s)	S. Pesant (EMBL-EBI)
Due date	31.03.2022
Submission date	31.07.2022
Type	Handbook (Report)
Status and version	Vesrion 2 of 4 to be delivered every year

Declaration: *Any work or result described therein is a genuine output of the AtlantECO project. Any other source will be properly referenced where and when relevant.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 862923. This output reflects only the author's view and the European Union cannot be held responsible for any use that may be made of the information contained therein.

Table of Content

1 Version History

Version	Authors	Summary of changes	Date
0.1	S. Pesant (EMBL-EBI)	D4.1 Handbook V1	24.05.2021
0.2	S. Pesant (EMBL-EBI)	Addition of sample collection protocols package	01.05.2022
0.3	E. Muxagata (FURG) & D. Johns (MBA)	Addition of CPR collection protocols package	01.05.2022
0.4	J. Poulain (CEA)	Addition of genetic analysis protocols package	01.06.2022
0.5	A. Elineau (SU)	Addition of imaging analysis protocols package	01.06.2022
0.6	M. Huete (EMBRC-UVIGO) J. Schramm (FTO)	Addition of ABS protocols package	01.07.2022
0.7	S. Pesant (EMBL-EBI)	D4.2 Handbook V2 text updated	30.07.2022

2 Executive summary

The present deliverable (D4.2) constitutes version 2 of the Handbook of Standards and Best Practices regarding AtlantECO's augmented observations and technological innovations as described in Work Package 4. The Handbook will be developed throughout the project and will be delivered progressively as four revised and augmented deliverables (D4.1-D4.4). The Handbook will address standards and best practices with respect to:

- Sample collection
- Sample biobanking
- Sample analysis
- Access and Benefits Sharing
- Provenance reporting
- Environmental context reporting

AtlantECO will generate new observations as part of the activities described in its Grant Agreement and also via a number of synergies with other European and International projects. The Handbook aims to harmonise as much as possible the methodologies used across these projects on some key types of observations, while allowing enough flexibility to adapt methodologies to the different observation programmes, therefore maximising the adoption of the standards and best practices. The present deliverable augments the Handbook with a collection of packages made publicly available on Zenodo (https://doi.org/10.5281/zenodo.4897860).

The Handbook includes from the previous version:

- A short introduction to the Project's goals, organisation and expected outcomes
- Guiding principles for the design of AtlantECO's system of standards & best practices

The system of base protocols presented in version 1 is augmented with two packages:

- The Mission Microbiomes handbook of sampling protocols
- The Continuous Plankton Recorder handbook of sampling protocols

The capacity building strategy presented in version 1 is augmented with three packages:

- A genetic analysis package
- An imaging analysis package
- An Access and Benefits Sharing package

This document is based on the terms and conditions established in the Grant Agreement (GA) and its Annexes, as well as in the Consortium Agreement (CA).

3 Introduction

The AtlantECO Project aims to study the Atlantic Ocean from pole to pole to determine the structure and function of the Atlantic microbiome in the context of ocean circulation and presence of pollutants to assess 1) its role in driving the dynamics of Atlantic ecosystems at basin and regional scales; 2) its potential to act as an indicator of ecosystem status and 3) the mechanism by which it drives the provision of ecosystem services.

In order to deliver this, the Project builds on four activity streams (AS):

- **AS1 - Asses the status of the ecosystem structures, functions, health and services** at regional, basin and all Atlantic scales and provide high quality gridded data products and maps;
- **AS2 - Enhance knowledge and innovate** by adopting standard optical and genetic observations protocols, cutting-edge network analysis methods and better parametrisation of connectivity and biogeochemical models;
- **AS3 - Assess drivers and stressors of change and forecast** their impact on tipping points and recovery of ecosystem structures, functions and services and develop eco-socio-economic models to predict their future states;
- **AS4 - Share and use** capacity and knowledge across the four continents bordering the Atlantic Ocean ensuring a seamless engagement between science, industry, policy and society.

AtlantECO is organised around three pillars of research: **marine microbiomes**, **plastics and the plastisphere** and **seascape and connectivity**, and focuses on five ecosystem services:

- **ES1**: Climate support system: carbon drawdown and storage
- **ES2**: Deep ocean life support system: transfer of matter & energy to mesopelagic and seabed ecosystems.
- **ES3**: Food security and the trophic fluxes: fisheries and aquaculture.
- **ES4**: Healthy planet, healthy people: biohazards, cycling of plastics and pollutants
- **ES5**: Biodiversity: resilience, adaptation and contribution to the Bioeconomy, including new chemicals

The Project addresses how the three pillars of research, together affect the ecology, biodiversity, sensitivity to climate change and the sustainability of the five ecosystem services. Building on this knowledge AtlantECO will optimize a set of tools and metrics into a unified framework of Eco-Socio-Economic analyses and projections that tightly couple ecosystem functioning and socio-economic activities.

As part of its second activity stream (AS2), AtlantECO will generate new observations about its three pillars of research. These observations will be obtained from a wide range of sampling platforms, including public and private research vessels (AtlantECO's flagship cruises), coastal and offshore observatories (AtlantECO's All Atlantic Ocean Sampling day), and citizen sail boats (AtlantECO's citizen Sail for Science). In addition, a number of synergies with other European and International projects will contribute a large quantity of observations.

AtlantECO aims to harmonise as much as possible the methodologies used across these activities on some key types of observations, while allowing enough flexibility to adapt these methodologies in the framework of different observation programmes, therefore maximising the adoption of standards and best practices. In order to reach that goal, AtlantECO will develop a Handbook describing standards and best practices regarding:

- Sample collection
- Sample biobanking
- Sample analysis
- Access and Benefits Sharing
- Provenance reporting
- Environmental context reporting

The present deliverable (D4.1) is the first version of the Handbook, with a focus on biological sample collection protocols. It proposes:

- Guiding principles for the design of AtlantECO's system of standards & best practices
- A package of protocols that can adapt to various field observation programmes
- A capacity building strategy that facilitates the adoption of the system by the community

The Handbook will be revised and augmented progressively during the course of the project. Community consultations will happen every year in the form of workshops, and via peer-reviewed contributions to a special issue co-edited by AtlantECO and the Marine Microbiome Forum in the journal Frontiers in Microbiology[1]. A first online workshop was held in November 2020 in order to scope out the various protocols used by the community[2] and a first webinar about the proposed system was held in May 2021.

[1] Marine Microbiomes: Towards Standard Methods and Best Practices https://www.frontiersin.org/research-topics/15877/marine-microbiomes-towards-standard-methods-and-best-practices#overview

[2] AtlantECO, Consortium. (2020, November). AtlantECO - Workshop on Standards and Best Practices, 2020 Edition. Zenodo. http://doi.org/10.5281/zenodo.4275504

4 Guiding principles

Funding bodies, the industry and the scientific community largely endorse the principle that best practices and common standards underpin reproducibility, ensuring that the relevant elements of a methodology are used consistently and generate results of the highest quality. Even though standards and best practices should stand on their own, they should also function well together and facilitate the integration of pre-existing protocols.

Best practices are generally defined in a bottom-up fashion, by a community for a community, and commonly take the form of community handbooks, cookbooks, guidelines, manuals and standard operating procedures (SOPs). While the emergence of best practices is essential to support innovation, the many best practices adopted across different communities are often fragmented, with gaps and duplication, which often limits their combined use. Over the next four year, AtlantECO's Handbook will include a compendium of community best practices that are relevant to its three pillars of research, highlighting commonalities and differences.

Standards on the other hand are generally defined in a top-down fashion by authoritative bodies such as the International Standards Organisation (ISO), and they may serve as important benchmarks to evaluate best practices. The ISO defines standards as "documents of requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose." Over the next four year, AtlantECO's Handbook will analyse the various community standards and propose standards. They will include for example the recommendation of specific pieces of equipment and consumables, guidelines regarding the preservation of samples, and the use of controlled vocabularies and standard units when reporting data and metadata.

As a starting point, AtlantECO proposes a coherent system upon which standards & best practices will be built. The system is inspired by the ***International System of Units***[3], and it is based on the pioneering work of the ***Tara Oceans expedition***[4] and the ***European project MicroB3***[5]. Such a pragmatic approach will allow AtlantECO to rapidly deploy a first suite of protocols across its activities in order to reach its goals, while working on the aggregation and harmonisation of additional, community specific best practices.

4.1 International System of Units

The International System of Units (SI, abbreviated from the French Système international d'unités) is the modern form of the metric system. It is the only system of measurement with an official status in nearly every country in the world. It comprises a coherent system of units of measurement starting with seven base units. The system allows for an unlimited number of additional units, called derived units, which can always be represented as combinations of the base units. Nine derived given as example in Figure 1.

The SI also provides twenty prefixes to the unit names and unit symbols that facilitate the nomenclature and semantic of the system. The SI is intended to be an evolving system; units and prefixes are created and unit definitions are modified through international agreement as the technology of measurement progresses and the precision of measurements improves.

[3] International Bureau of Weights and Measures (2006), The International System of Units (SI), ISBN 92-822-2213-6

[4] Karsenti et al. (2011) A Holistic Approach to Marine Eco-Systems Biology. PLoS Biol 9(10): e1001177. https://doi.org/10.1371/journal.pbio.1001177

[5] Ten Hoopen et al. (2015) Marine microbial biodiversity, bioinformatics and biotechnology (M2B3) data reporting and service standards. Stand Genomic Sci. 10(20). https://dx.doi.org/10.1186%2Fs40793-015-0001-5

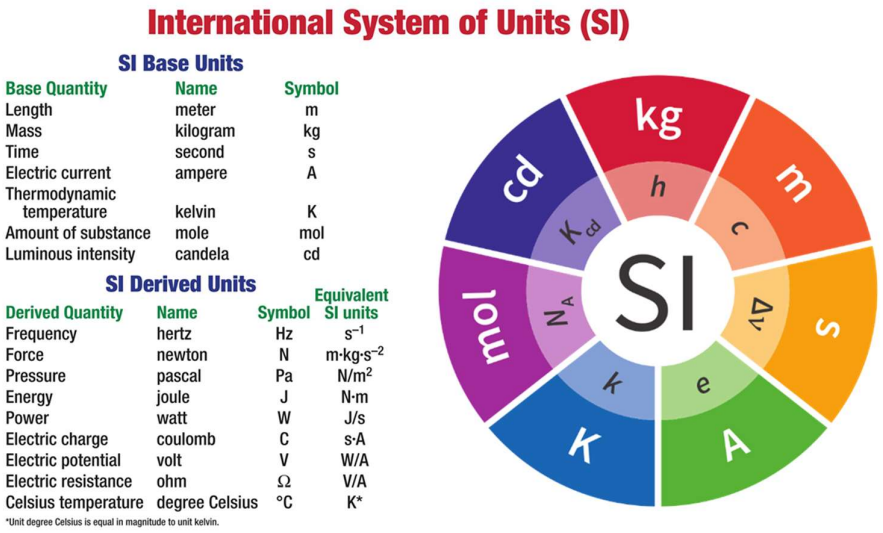


Figure 1. The seven base units and examples of derived units from the International System of units.

4.2 Tara Oceans end-to-end protocols[6]

The aim of Tara Oceans was to sample most of the global ocean's biogeographic provinces and to follow standardised protocols in order to facilitate comparative analysis. Seawater was sampled from the sunlit, epipelagic zone and from the dark, mesopelagic zone using water bottles, pumps and nets down to 1,000 m. Plankton were separated according to their size using sieves and filtering membranes, and were preserved in liquid nitrogen or with different fixatives for molecular and/or morphological analyses. Back in the laboratory, nucleic acids were extracted from the filters and subjected to high-throughput sequencing (HTS) to generate metabarcoding (metaB), metagenomic (metaG) and metatranscriptomic (metaT) data sets as well as to yield single-cell genomes. Deep sequencing was performed with Illumina technology at high coverage rates per sample to access the genomic content of plankton species, including those that are rare in the environment. Other samples were analysed with high-throughput imaging (HTI) methods to determine the taxonomy of plankton across size fractions spanning seven orders of magnitude, and to quantify their abundance, cellular biovolumes and other morphological attributes.

High-throughput sequencing (HTS) and imaging (HTS) are complementary, but they are rarely sampled concurrently and combined for analysis. Together, they offer powerful means to reveal how environmental and/or genomic variability impacts organismal or cellular phenotypes, and to generate hypotheses about the nature of physical interactions between organisms.

[6] Sunagawa et al. (2020) Tara Oceans: towards global ocean ecosystems biology. Nat Rev Microbiol 18, 428–445. https://doi.org/10.1038/s41579-020-0364-5

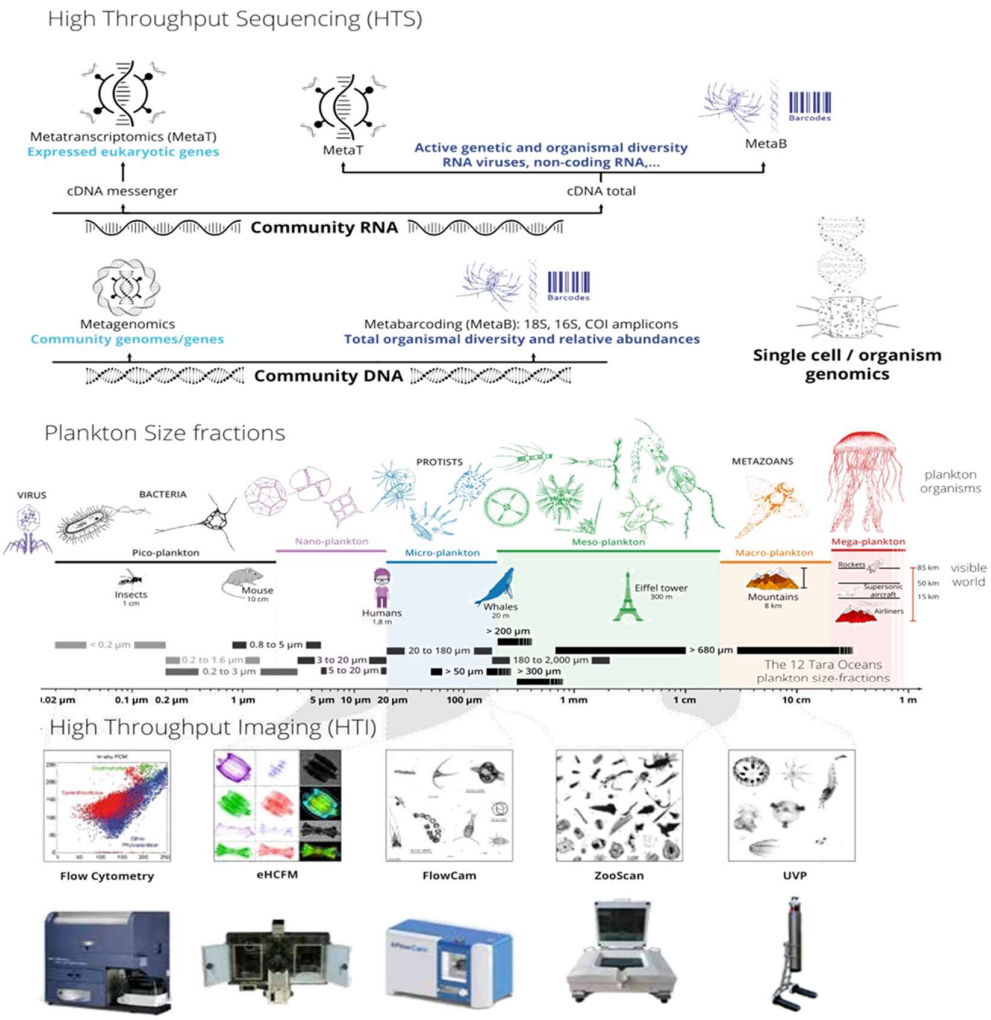


Figure 2. The base entities (middle part), and a range of sequencing analysis methods (top) and imaging analysis methods (bottom) targeted by the sampling strategies of Tara Oceans (reproduced from Suanagawa et al. 2020).

High Throughput Sequencing (HTS) included metabarcoding (metaB) sequencing to provide a baseline survey of the diversity and relative abundance of prokaryotic and eukaryotic taxa in their environmental context. n addition, metagenomic (metaG) and metatranscriptomic (metaT) sequencing were performed to uncover the genomic potential and expression of genes in viruses, prokaryotes, and eukaryotes, with the additional advantage of gaining insights into the population structure and evolution of, and selective pressure on, the most abundant planktonic organisms in the ocean. Finally, additional efforts went into the sequencing of new draft genomes from single cells of protists and bacteria that were isolated by flow cytometry from cyropreserved seawater samples, as well as of individual zooplankton organisms.

High Throughput Imaging (HTI) was achieved using a series of automated imaging methods to quantify the concentration, biomass, biovolume, taxonomic composition, and morphological features of plankton, as well as suspended and sinking particles across organismal size fractions. These methods included the Underwater Vision Profiler, ZooScan, FlowCAM, Imaging FlowCytobot, and e-HCFM, an environmental high-content fluorescence microscopy workflow developed during the project. Together, these technologies allowed for automated imaging of organisms across different taxonomic and functional groups ranging in size from individual cells to large gelatinous zooplankton and marine snow particles.

5 Building a system of base protocols and derived protocols

AtlantECO proposes a system that comprises seven bases of protocols for the sequencing and imaging of the Ocean Microbiome, including that of relatively large multicellular eukaryotic organisms such as zooplankton. The seven bases partition the Ocean Microbiome into a continuum of six groups of biological entities, plus one "environmental" base that conveniently encompasses protocols that fall outside the framework of the other six bases. Each base targets a group of biological entities according to their organisation at molecular, cellular and community scales, and according to their size.

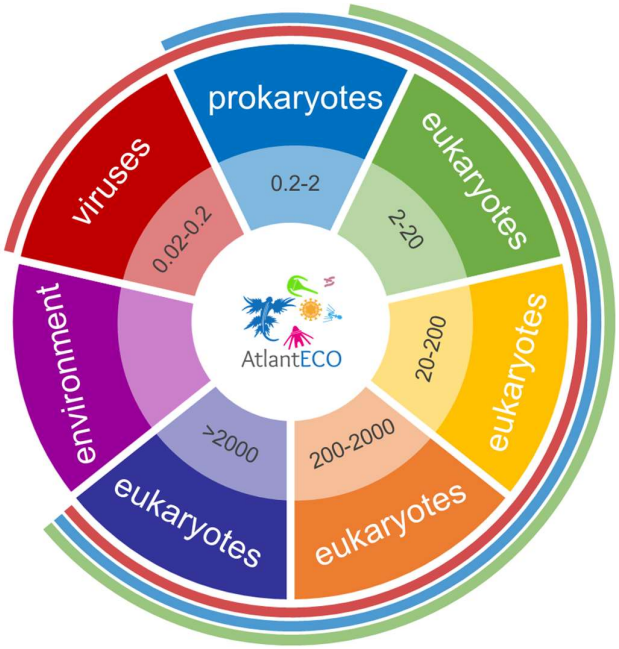


Figure 3. Schematics of AtlantECO's system of seven base protocols for the sequencing and imaging of Ocean Microbiomes, including large eukaryotic plankton.

The first six bases are targeting:

1. ***viruses*** in the size range 0.02-0.2 µm, which lack a nucleus and bear only RNA or DNA, but not both;
2. ***prokaryotes*** in the size range 0.2-2 µm, which lack a nucleus but bear both DNA and RNA;
3. ***eukaryotes*** in the size range 2-20 µm, which have a nucleus and bear both DNA and RNA, and essentially comprise unicellular organisms;
4. ***eukaryotes*** in the size range 20-200 µm, which have a nucleus and bear both DNA and RNA, and comprise unicellular, colonial and some multicellular organisms
5. ***eukaryotes*** in the size range 200-2000 µm, which have a nucleus and bear both DNA and RNA, and comprise essentially multicellular and occasionally large or colonial unicellular organisms; and
6. ***eukaryotes*** in the size range >2000 µm, which have a nucleus and bear both DNA and RNA, and comprise the occasional large organisms caught in plankton nets, e.g. salps, jellyfish and fish larvae.

The complexity of biological entities cannot be simply resolved by finite partitions, and as a result, entities and processes can span several of the proposed partitions. This is depicted by three outer rings on figure 3, illustrating that biological entities from the three smaller size fractions are often present in other partitions as a result of biological associations such as symbiosis, parasitism, aggregation, or sometimes as a result of

methodological biases. The significance and implications of this complexity will be addressed by the proposed system. The seventh base of protocols will include imaging protocols such as remote sensing algorithms, and sequencing protocols such as eDNA, which aims to be all encompassing and may target organisms outside the realm of the Ocean Microbiome such as fish, reptiles, birds and mammals.

General principles behind the partitioning of base protocols:

- *Base protocols are partitioned according to the organisation of biological entities, considering molecular, cellular and community scales.*
- *Base protocols are partitioned according to the size of biological entities, encompassing the end-to-end, continuous size spectrum of the Ocean Microbiome*
- *The complexity of biological entities cannot be simply resolved by finite partitions, and as a result, entities and processes can span several of the proposed partitions.*

AtlantECO's system of base protocols uses simple sets of prefixes and suffixes that facilitate the nomenclature and semantic of the system. The label of the base protocols is composed of one letter designating the target analysis, i.e. the letter "S" for sequencing or "I" for imaging, followed by one to several digits designating the lower threshold of the size fraction (Figure 4). Even though, in practice, the sampling gear and the membranes used to partition biological entities may vary in mesh or pore size, the standard label of the base protocols is generic. For example, membranes with pores size 3 and nets with mesh size 5 µm are often used interchangeably to partition the second and third base protocols, but the label of the third base sequencing protocol is consistently "S2" and its target size fraction is consistently "2-20 µm".

Each base protocol comes with a set of targets regarding, for example, biological entities, size fractionation, the required volume of seawater, processing time, and preservation methods (Figure 4). The Handbook will recommend specific equipment, consumables and best practices to reach the various targets. However, standards may take years to be adopted by the community, and we recognise that for a number of good reasons such as lack of capacity or consistency with historical practices, community best practices may differ from those targets. Such cases will constitute "derived protocols". The system of base protocols aims to provide benchmarks against which various methodologies and best practices can be assessed, and as much as possible to provide an intercomparison between derived and base protocols. In all cases, the specificity of derived protocols will be documented in the provenance metadata of the samples, according to reporting standards that will be defined in version 2 of the Handbook.

The Handbook will grow significantly during the four years of the project as it addresses topics such as (1) Sample collection, (2) Sample biobanking, (3) Sample analysis, (4) Access and Benefits Sharing, (5) Provenance metadata reporting, and (6) Environmental contextual data reporting. The recommended standards and best practices may also evolve rapidly with the continuing progress in accuracy, throughput, and cost reduction of novel methods such as the long-read sequencing of native DNA and RNA, which eliminates the amplification bias of short-read sequencing and opens a broad range of applications for Ocean Microbiome genomics.

Base sequencing protocols

Base Protocol Label	Target Entity	Target Size fraction (µm)	Target Analysis	Target Volume (L)	Target Time (min)	Target Preservation
S002	Virus	0.02-0.2	MetaG, MetaT	20	Flocculated 4-24h	+4°C
S02	Prokaryotes	0.2-2	MetaB, MetaG, MetaT	20	<15	LN_2 or -80°C
S2	Eukaryotes	2-20	MetaB, MetaG, MetaT	20	<15	LN_2 or -80°C
S20	Eukaryotes	20-200	MetaB, MetaG, MetaT	10^2-10^4	<15	LN_2 or -80°C
S200	Eukaryotes	200-2000	MetaB, MetaG, MetaT	10^3-10^5	<15	LN_2 or -80°C
S2000	Eukaryotes	>2000	MetaB	10^3-10^5	<15	ethanol
S0	Environment	unfractionated	eDNA	2		buffer

Base imaging protocols

Base Protocol Label	Target Entity	Target Size fraction (µm)	Target Analysis	Target Volume (L)	Target Time (min)	Target Preservation
I002	Virus	0.02-0.2	Flow cytometry	10^{-3}	na	LN_2 or -80°C
I02	Prokaryotes	0.2-2	Flow cytometry	10^{-3}	na	LN_2 or -80°C
I2	Eukaryotes	2-20	Fluorescence microscopy	10^2	na	+4°C
I20	Eukaryotes	20-200	Flow imaging microscopy	10^2-10^4	na	live
I200	Eukaryotes	200-2000	Flatbed scan imaging	10^3-10^5	na	formaldehyde
I2000	Eukaryotes	>2000	Light microscopy	10^3-10^5	na	ethanol
I0	Environment	>2 km	Remote sensing	na	na	na

Figure 4. Summary of AtlantECO's base protocols (version 1) for the sequencing and imaging of Ocean Microbiomes

5.1 The Mission Microbiomes handbook of sampling protocols

As part of Mission Microbiomes, AtlantECO developed a handbook of protocols for the collection of samples that target the biogeochemical, genetic and imaging analysis of marine microbiomes, plastics and the plastisphere. It is available on Zenodo, together with the present deliverable (https://doi.org/10.5281/zenodo.4897860). Although the handbook provides specific instructions for the implementation of protocols on board AtlantECO's flagship *SV Tara*, it introduces generic tools and best practices that are used across AtlantECO's sampling activities, i.e. other flagship cruises, the All Atlantic Ocean Sampling Day, the Citizen Sail-4-Science, the CPR Brazil-South-Africa line, and third party cruises. The generic tools include sample labels (stickers) with a unique identifier, logsheets that capture standard metadata about the provenance and the methodological processing of samples, flowcharts that summarise the preparation of samples, and detailed step-by-step protocols.

Unique sample identifiers, in the form of an alpha-numeric string and its barcode equivalent, are pre-printed in duplicate on a pair of label stickers. One of the labels is small with only the barcode printed on it, and the other one is larger, with a barcode and enough space to write with a pen. The large label is affixed on the sample container (e.g. tube, bottle, whirlpack, petri-slides), whereas the small one goes on the corresponding logsheet. The "wrapping" labels are designed for 2-mL, 5-mL and 15-mL cryotubes stored in liquid nitrogen of in a freezer. The "normal" labels have a second small label that is sometimes used for example on the cap of larger sample container.

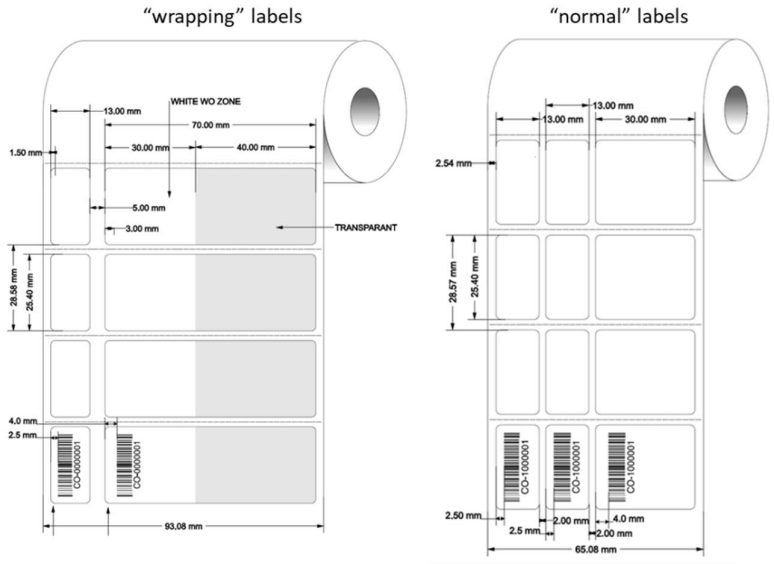


Figure 5. Sample labels come in pairs, pre-printed with the same unique barcode identifier.

Most importantly, labels are pre-printed with unique identifiers that are pre-registered at BioSamples (https://www.ebi.ac.uk/biosamples/), so that the same identifier is used on the sample itself, on printed and digital metadata records generated during sampling, in data files back in the analysis labs, and in the data archives. Provenance metadata and environmental context are uploaded on BioSamples, which are then linked to images archived in BioImage (https://www.ebi.ac.uk/bioimage-archive/), sequences archived in ENA/NCBI/BBNJ (https://www.ebi.ac.uk/ena/), proteomes archived in PRIDE (https://www.ebi.ac.uk/pride/), metabolomes archived in Metabolights (https://www.ebi.ac.uk/metabolights/), and to taxonomic and functional annotations available at MGnify (https://www.ebi.ac.uk/metagenomics/).

Paper logsheets are printed on both sides, typically with predetermined fields on the front to (1) capture georefereces at the start and end of the sampling event, (2) affix sample labels on a "BINGO" style grid where columns are protocols and rows are either distinct sampling depths, replicates, or controls, (3) write processing times and volumes. The back of the logsheet provides labelling instructions and space for additional comments. Each logsheet is scanned and stored as PDF files in a public archive (e.g. https://store.pangaea.de/Projects/TARA-MISSION-MICROBIOMES/Logsheets/), so that the original metadata can be visualised. Sample unique identifiers, i.e. barcodes printed on labels, are read automatically using a software, and the position of the barcode on the pdf provides information regarding the corresponding protocol and sampling depth/replicate/control. Automatic recognition of hand-written information such as the provenance metadata and the processing time and volume is being developed, so we currently gather the information by manual curation.

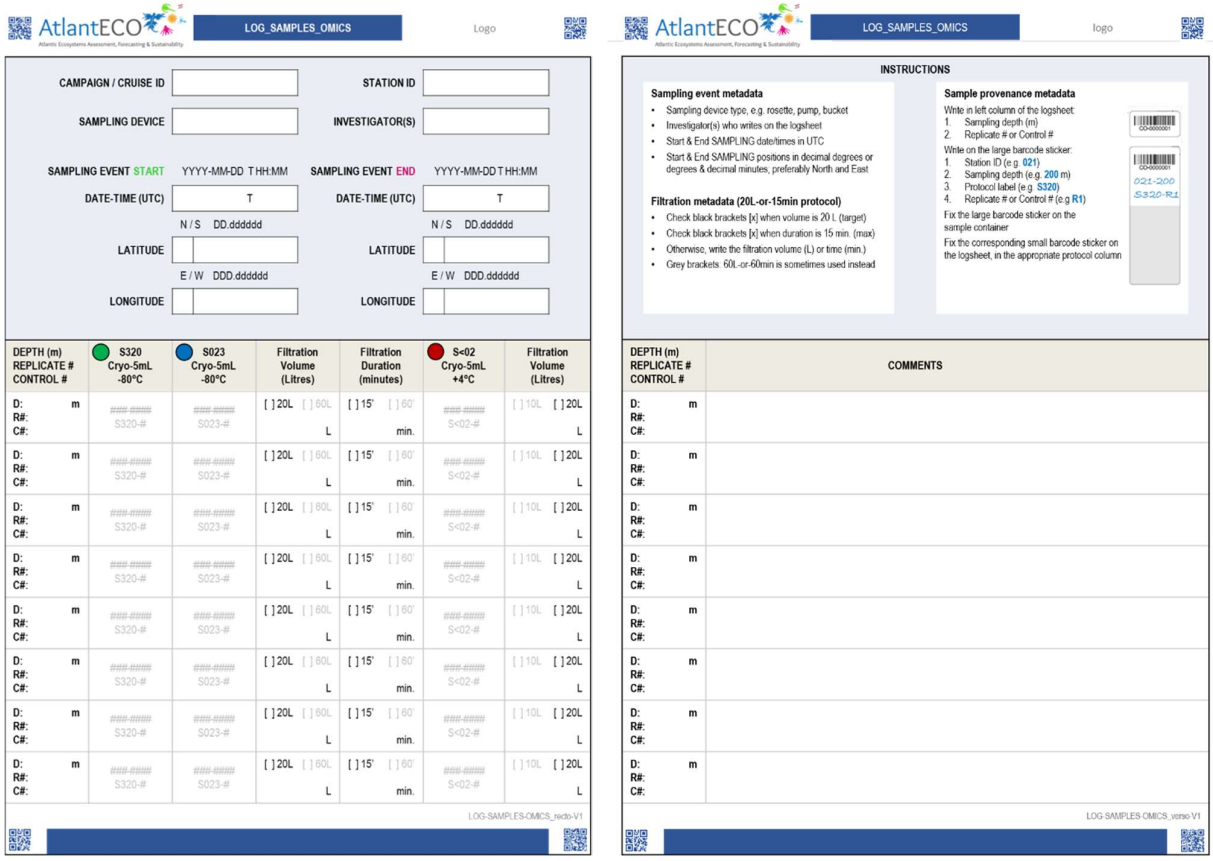


Figure 6. Paper logsheets are printed on both sides, typically with predetermined fields on the front, and instructions and comments on the back.

Over 60 protocols are included in the Handbook, including the 13 base protocols targeting the genetic and imaging analysis of marine microbiomes (Figure 5). It also includes biogeochemical protocols used to provide environmental context, protocols for the genetic and imaging analysis of plastic & the plastisphere, and two suites of size-fractionated protocols targeting the proteomic and metabolomic analysis of the ocean microbiome. These are summarised on flowcharts and step-by-step protocols.

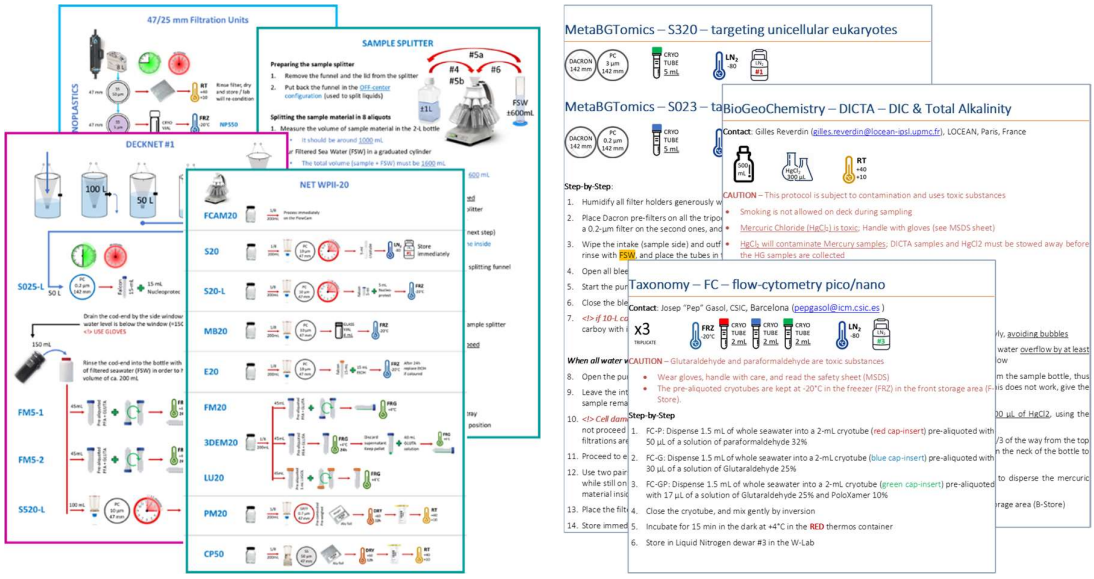


Figure 7. Protocol flowcharts (left) and step-by-step description provided in the Mission Microbiomes handbook.

5.2 The Continuous Plankton Recorder handbook of sampling protocols

The continuous plankton recorder survey is the most geographically extensive marine monitoring programme in the world. Yet, large areas of the world's oceans do not have regular CPR surveys, notably the sub-tropical and tropical regions of the Atlantic where AtlantECO has launched the first regular CPR line between Brazil and South Africa (Figure 8). The sampling and sample analysis protocols are available on Zenodo, together with the present deliverable (https://doi.org/10.5281/zenodo.4897860).

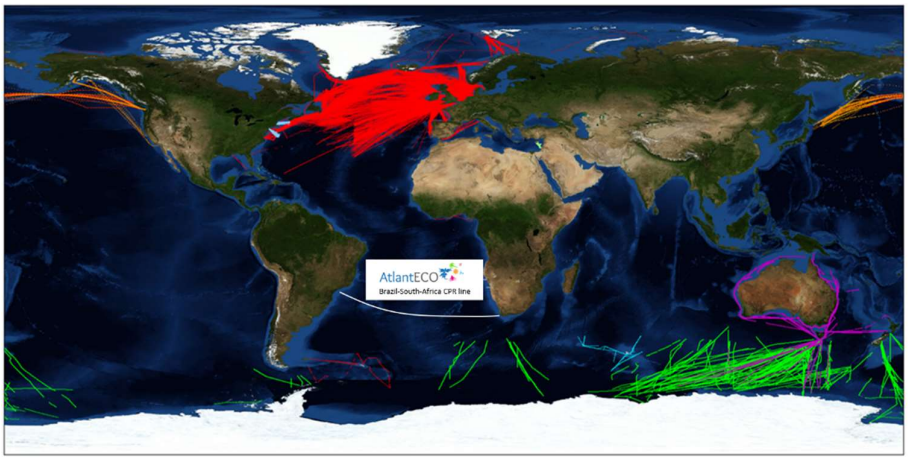


Figure 8. AtlantECO's CPR line across the South Atlantic, a scientific collaboration between Brazil and South Africa.

6 Building capacity for the adoption of the system by the community

The value of a system can be measured by the degree to which it is adopted by its users. This will be achieved in part by involving the community in defining the system in a coordinated effort with UNESCO's Ocean Best Practices System (OBPS)[7] where AtlantECO will contribute to, and benefit from their online resources as well as their outreach and engagement activities. In the present section we explain how the proposed base protocols can be used as "plug and play" components of a sampling strategy, thus improving their adoption in a wide range of use cases and demonstrating their fitness for purpose. We also explain how we plan to build capacity around the system of base protocols, notably with training and the development of sampling kits.

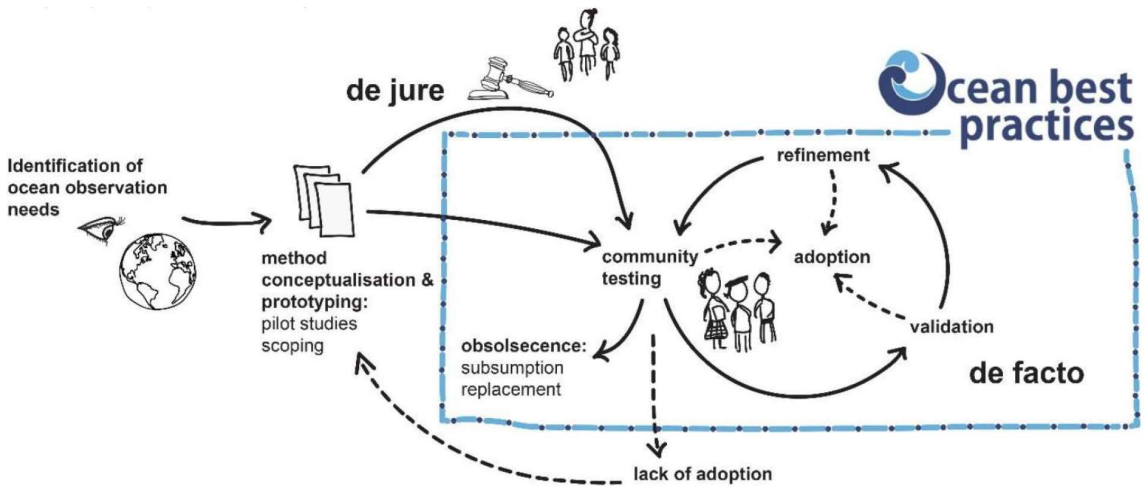


Figure 9. The framework proposed by OPBS for the adoption of best practices[8]

The adoption of AtlantECO standards and best practices by the community depends on the capacity of a wide variety of stakeholders to carry out the proposed protocols, including small marine stations, large institutions, national and international sampling programmes, industry, and ultimately citizen scientists. Ahead of AtlantECO's pilot All Atlantic Ocean Microbiome Sampling, we conducted a short survey about community best practices and capacity to carry out base protocols or derived protocols for the sequencing of Ocean Microbiomes in the size fractions targeting prokaryotes and eukaryotes (S02 & S2; Figure 4). The survey was taken by an extended list of candidate sampling sites, especially along the Atlantic coast of Africa. Asked about their readiness to take part in the pilot All Atlantic Ocean Microbiome Sampling, 1/3 are ready to sample on the 21st of September 2022, whereas nearly 2/3 are ready to sample between the 21st of September and the 21st of December 2022 (Figure 10).

[7] https://www.oceanbestpractices.org/

[8] Hörstmann et al. (2020) Towards a Best Practice for Developing Best Practices in Ocean Observation (BP4BP). IOC Manuals and Guides No.84 (IOC/2020/MG/84). http://dx.doi.org/10.25607/OBP-781

Figure 10. Readiness of the All-Atlantic community to participate in the pilot All Atlantic Ocean Microbiome Sampling.

The survey results show that the best practice developed several decades ago for environmental monitoring of phytoplankton is still mostly used across the community. The best practice can be summarised as processing relatively small volumes of seawater on relatively small size filters, using a vacuum filtration system. This best practice reportedly results in relatively long filtration times and the generation of a single size fraction. The combination of filtrations in series or in parallel using a sterivex and small-to-large diameter (47-142 mm) filters is also adopted variably by the community. The methods used to preserve samples are also variable and we are now considering adopting the use of DNA/RNA shield to preserve material while being shipped at room temperature. Although the proposed base protocol (A) is currently not a widely adopted best practice, the survey results show that many have capacity to carry it out, and many others wish to build such capacity. These results are extremely encouraging for AtlantECO's ambition to augment the capacity to sample the Ocean Microbiomes along and across the Atlantic.

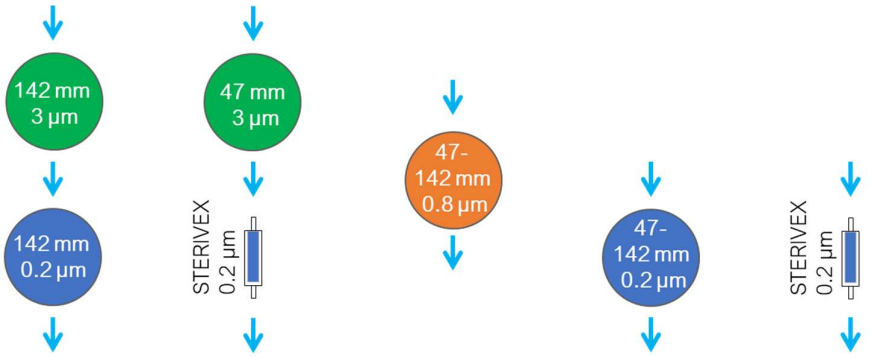
AtlantECO partners are working on several kits and innovative technologies that will build capacity around the proposed standards. The Handbook introduces two kits. The "filtration kit" will greatly improve the adoption of the first three base protocols for sequencing (S002, S02 and S2), whereas the "planktoscope" will undoubtedly boost the uptake of the fourth base protocol for Imaging (I20) by academia, industry and citizen science communities.

The production and distribution of sampling kits inevitably incurs a cost. While AtlantECO can cover in part the costs of kits for its flagship Mission Microbiomes, significant additional funding is needed in order to roll-out sampling kits in marine stations bordering the Atlantic Ocean and on board international oceanographic sampling programmes. In the context of the Belem agreement, AtlantECO partners are looking for co-funding opportunities in Europe, South Africa and Brazil. In the context of the Galway agreement, coordinated funding opportunities with NOAA in the United States and Fisheries and Oceans Canada will be facilitated by the AORA Marine Microbiome Working Group[9]. Efforts to distribute kits in other countries bordering the Atlantic will also be initiated, particularly in countries that have bilateral research agreements with the European Union, such as Morocco, Argentina and Cape Verde. Finally, in the context of the Decade of Ocean Science for Sustainable Development (2021-2030), AtlantECO will seek endorsement and support from the UNESCO field offices, networks, and institutes for capacity building.

The adoption and use of sampling kits by communities will require training activities. Detailed written protocols, as well as short, pre-recorded video training and webinars will be provided with the Handbook.

[9] https://www.marinemicrobiome.org/

Together with the European Marine Biology Resource Centre (EMBRC), AtlantECO has organise a hand-on training sessions with the kits during a 2022 summer school in Cape Town. It is anticipated that similar training sessions will be included in summer schools organised in synergy with other European and International Initiatives.



METHODS	A	B	C	D	E
Filters	2x 142 mm membranes	47 mm membr. + sterivex	47 or 142 mm membrane	47 or 142 mm membrane	sterivex
Vol. in 15 min.	10-20 L	2-5 L	1-10L	1-10L	1-10 L
Pump system	large peristaltic	small peristaltic	various	various	various
ADOPTED BY	A	B	C	D	E
Ocean Sampling Day			(✓)		✓
Bio-GO-SHIP*					✓
Tara Oceans	✓		(✓)	(✓)	
EMO-BON**	✓				

*Bio-GO-SHIP Linking marine biodiversity and biogeochemistry (https://biogoship.org/)
**European Marine Omics Biodiversity Observatory Network (https://www.embrc.eu/emo-bon)

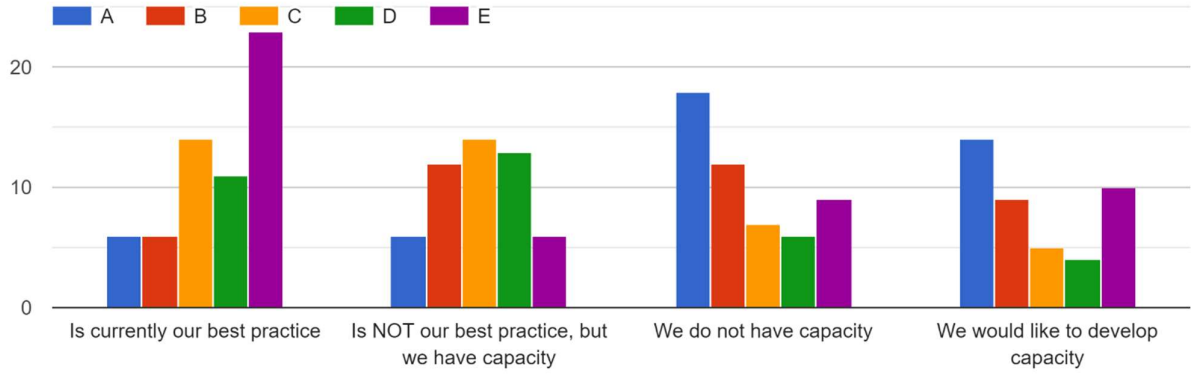


Figure 11. Results of the survey on community best practices and capacity to carry out base protocols or derived protocols for the sequencing of Ocean Microbiomes in the size fractions targeting prokaryotes and eukaryotes.

6.1 Filtration kit for key sequencing protocols

With the ambition to analyse metabarcodes, metagenomes and metatranscriptomes of viruses, prokaryotes and small size eukaryotes comes the requirement to filter tens of litres of seawater in a relatively short time. It also requires the fractionation of biological entities according to size, so that the genetic material of less abundant, larger organisms is not shadowed, as a result of sequencing, by that of relatively more abundant, smaller organisms.

These requirements call for using robust benchtop peristaltic pumps and large filter holders. The operation of such equipment on board small and large research platforms under various weather conditions and in low- to high-biomass environments has its challenges, some of which affect the reproducibility of the method and the quality of the sample. Perhaps the most crucial steps of this operation are setting up membranes in the filter holders before the filtration starts and recovering them at the end. The latter step involves draining the filter holders, opening them, folding the filters and packaging them carefully. Most importantly, that step involves one brain, two hands and a great capacity to adapt and solve problems, which is perhaps our best asset as scientists, but can sometimes impair our efforts towards standardisation and reproducibility.

AtlantECO is developing a filtration kit that will improve both the reproducibility of these sampling protocols, and their adoption along and across the Atlantic. The kit will initially target the package of base sequencing protocols proposed in the previous section, i.e. consisting of S002, S02 and S2. Several kits are currently used on board AtlantECO's flagships Tara, Veleiro ECO, Alpha Crucis and Agulhas II.

The kit comprises commonly used components that have proven to be effective and reliable in the field, notably a Masterflex I/P series peristaltic pump, Masterflex silicone tubing (ID=9.5 mm), two Millipore 142-mm-diameter stainless steel filter holders, a manual vacuum pump, and two 20-litre Nalgene carboys. The kit will most importantly include a "filtration box" that facilitates and standardises the use of the Millipore filter holders (Figure 12). It will provide the following features and advantages:

- The two filter holders are securely fixed on the box, removing the clutter of the 6 "legs" and stabilising the overall setup
- The tubing connectors on each of the top sections are replaced by stainless steel quick-release connectors, thus facilitating the plug and play setup
- The quick-release connectors also free the top sections from their tubing once the filtration is over, which facilitates the drainage step
- Two slots in the box can hold the top section of each filtration head, freeing space, reducing risks of contamination, and helping to remove the filter from the bottom section of the filtration heads
- The tubing connectors on each of the bottom sections are replaced by elbow connectors, thus reducing the overall height and stability of the setup
- The tubes that carry the filtrate of the two filtration heads are inside the box and come out at the front of the box, avoiding the clutter of tubes, and allowing to easily collect the filtrate if needed
- The two filtration heads can be used individually for parallel filtrations, or can be connected together for serial filtrations
- The box can be setup with the two heads side by side or one in front of the other, adapting to the benchtop space and configuration
- The current dimensions of the prototype box are 50 cm x 30 cm x 20 cm (length x width x height). We hope to slightly reduce this footprint.

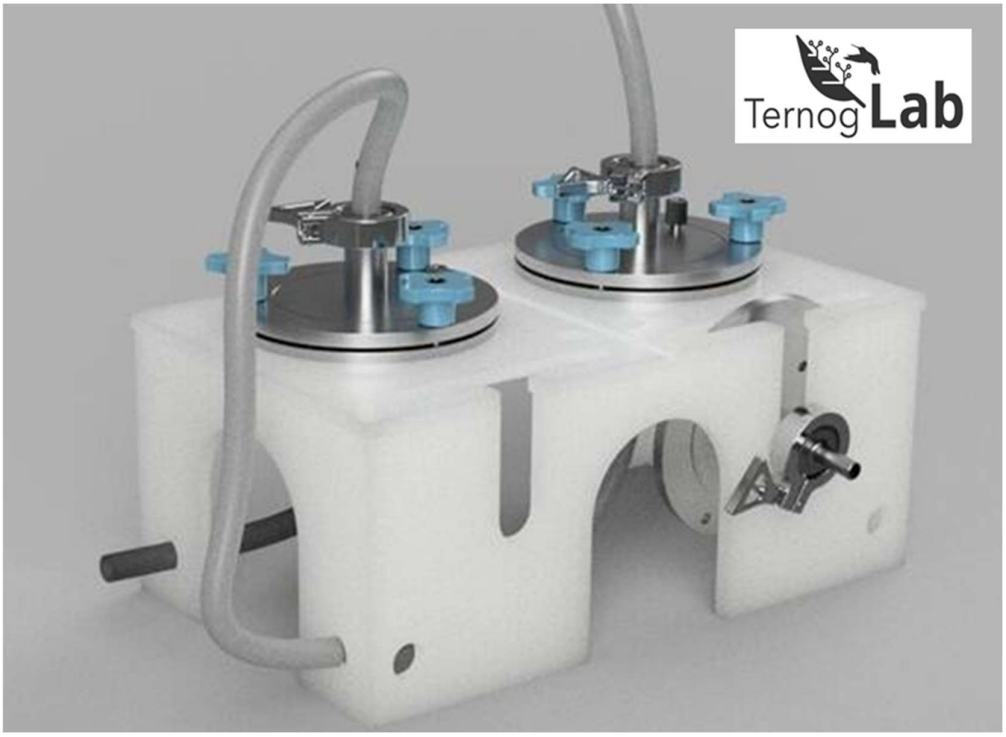


Figure 12. 3D model of the prototype "filtration box", without the quick-release connectors.

6.2 Sampling kit for key imaging protocols

Quantitative flow imaging microscopy is AtlantECO's fourth base protocol (I20) for the analysis of the Ocean Microbiome in the size range 20-200 µm. Several high-throughput and automated imaging instruments such as the Imaging FlowCytobot from McLane Research Laboratories Inc. and the FlowCam from Yokogawa Fluid Imaging Technologies Inc. have been developed to target this protocol. However, these instruments are expensive and relatively large, which may amper their wider adoption by marine stations, small research vessels, and ultimately citizen science programmes.

Technological innovations in the field of communication devices have led to the development of low cost, high performance and compact lenses and cameras. In collaboration with AtlantECO partner Sorbonne Université, Plankton Planet[10] developed the ***Planktoscope***, a high-performance, low-cost flow imaging microscopy instrument that combines these optical components with a small, dual-display, Raspberry PI4 computer and a small, off the shelf peristaltic pump (Figure 13). It is designed for robust usage and portability, and it achieves an optical magnification of 0.75X with a pixel size of 1.5 µm/px. The travel distance of the focus stage is about 3.2 cm for a step size of 0.15 µm. With a flow rate of ca. 3 ml/min and a flow-cell with a thickness of 500 µm, the instrument can image a volume of 1.7ml/min. The platform is operated via an open-software strategy, using existing libraries and a flow-based visual programming platform that enables any user to rapidly customize acquisition and processing steps. More specifically, it used python libraries such as Node-RED for

[10] https://planktonplanet.org/

the Graphical User Interface and first layer of programming interface, MorphoCut for handling the image processing from the raw images to the online platform, and EcoTaxa for classification and annotation.

Benchmarking exercises that compare the results of the planktoscope with that of the commercial FlowCam instrument are extremely encouraging[11], and additional tests are being organised with AtlantECO's imaging nodes in South America (USP), South Africa (NMU) and Europe (SU). The successful adoption of the fourth imaging protocol (I20) will require training a generation of young taxonomists with the online image annotation platform EcoTaxa[12]. This will be achieved in part through co-funding with a FFEM[13] initiative lead by AtlantECO's communication partner FTO.

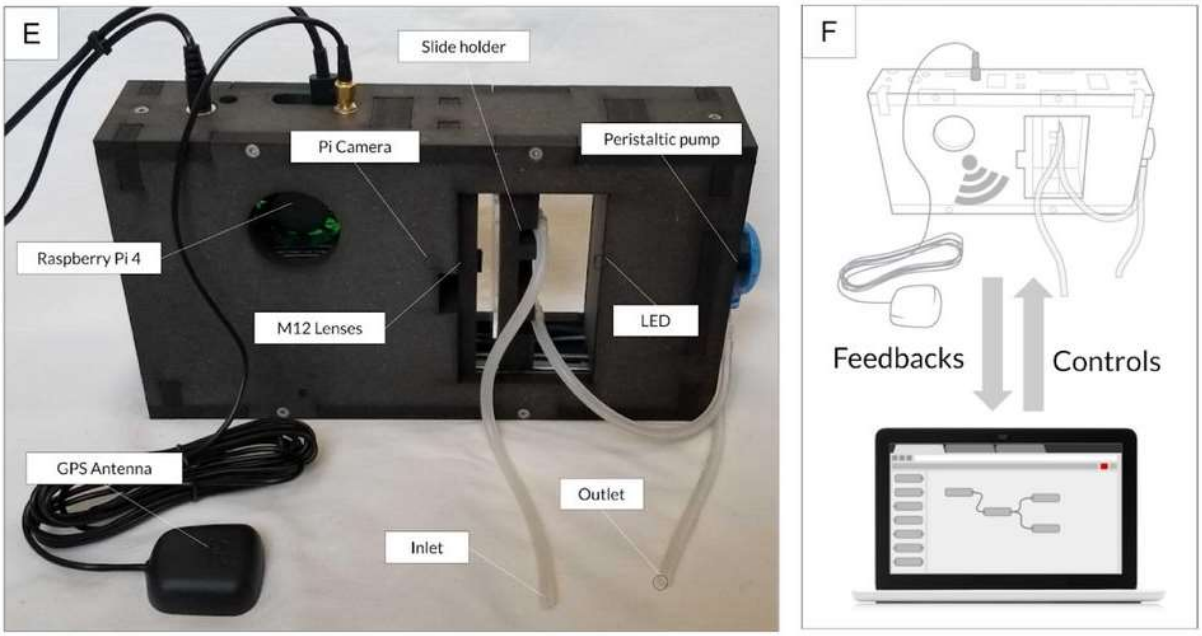


Figure 13. The planktoscope design and components.

6.3 Sequencing analysis protocols

AtlantECO adopted a "Hub & Nodes" approach in order to share analysis capacity and efforts among selected sequencing centres in Europe, Brazil and South Africa. Partner CEA acts as the Hub and hence provided reference protocols that will be used for an intercalibration exercise among Nodes, and the basis for the standard protocols that will be adopted for the sequencing of AtlantECO's augmented observations. The detailed analysis protocols are available on Zenodo, together with the present deliverable (https://doi.org/10.5281/zenodo.4897860).

[11] https://www.biorxiv.org/content/10.1101/2020.04.23.056978v1.full
[12] https://ecotaxa.obs-vlfr.fr/
[13] Fonds Français pour l'Environnement Mondial https://www.ffem.fr/fr/carte-des-projets/plancton-oceanique-climat-et-developpement

6.4 Imaging analysis protocols

Likewise, AtlantECO adopted a "Hub & Nodes" approach in order to share analysis capacity and efforts among selected imaging centres in Europe, Brazil and South Africa. Partner SU acts as the Hub and hence provided reference protocols that will be used for an intercalibration exercise among Nodes, and the basis for the standard protocols that will be adopted for the imaging of AtlantECO's augmented observations. The detailed analysis protocols are available on Zenodo (https://doi.org/10.5281/zenodo.4897860).

6.5 Access & Benefits Sharing protocols

AtlantECO partner EMBRC (UVIGO) and the European Blue Biobank (EBB) initiative have produced two guides on Access and Benefit Sharing (ABS) that we adopt as best practice. They are made available on Zenodo, together with the present deliverable (https://doi.org/10.5281/zenodo.4897860). The guides are complemented by a series of EBB Webinar on ABS that AtlantECO uses as part of its training activities.

We conducted a short survey about community awareness and capacity to meet the requirements of the United Nations Convention on the Law Of the Sea (UNCLOS), and the Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilisation. Regarding sampling permits, 38% of the community already has a recurrent sampling permit, 16% know the process to obtain one, and 28% are not familiar with the process of obtaining a sampling permit but will follow the EBB guidance to start the process. Regarding ABS, 46% of the community is familiar with the process of applying for Prior Informed Consent (PIC) that grants access to a genetic resource, whereas 42% are not but will follow the EBB guidance to start the process.

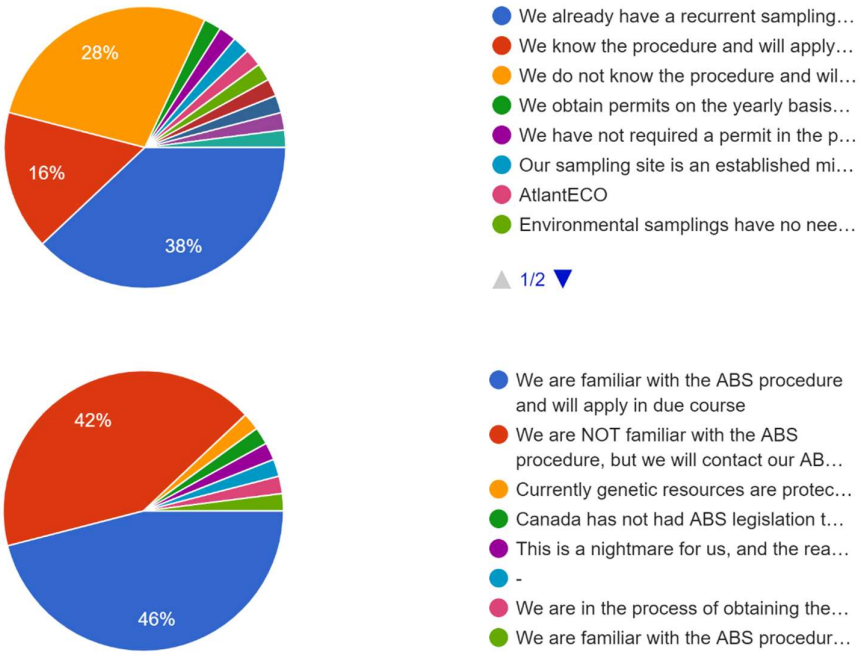


Figure 14. Community awareness and capacity to meet the requirements of the United Nations Convention on the Law Of the Sea (UNCLOS), and the Nagoya Protocol on access to genetic resource.