Project information	
Project full title	EuroSea: Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans
Project acronym	EuroSea
Grant agreement number	862626
Project start date and duration	1 November 2019, 50 months
Project website	https://www.eurosea.eu

Deliverable information	
Deliverable number	D3.19
Deliverable title	Omics community protocols
Description	List of Standard Operating Procedures (SOPs) for the integration of -omics protocols into marine observatories
Work Package number	WP3
Work Package title	Network Integration and Improvement
Lead beneficiary	SZN
Lead authors	Daniele Iudicone, Marina Montresor
Contributors	Raffaella Casotti, Lucia Campese, Anna Chiara Trano
Due date	31.08.2023
Submission date	16.10.2023
Comments	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 862626.

https://doi.org/10.3289/eurosea_d3.19

Table of contents

Executive summary

The aim of the WP3 “Network Integration and Improvements” is to coordinate and enhance key aspects of integration of European observing technology (and related data flows) for its use in the context of international ocean monitoring activities. One of the dimensions of the integrations is the constitution of thematic networks, that is, networks whose aim is to address specific observational challenges and thus to favor innovation, innovation that will ultimately support the Blue economy.

In this context, the specific aim of Task 3.8 is to accelerate the adoption of molecular methods such as genomic, transcriptomic (and related “omics”) approaches, currently used as monitoring tools in human health, to the assessment of the state and change of marine ecosystems. It was designed to favor the increase the capacity to evaluate biological diversity and the organismal metabolic states in different environmental conditions by the development of “augmented observatories”, utilizing state-of-art methodologies in genomic-enabled research at multidisciplinary observatories at well-established marine LTERs, with main focus on a mature oceanographic observatory in Naples, NEREA. In addition, an effort is dedicated to connecting existing observatories that intend to augment their observations with molecular tools.

Molecular approaches come with many different options for the protocols (size fractioning, sample collection and storage, sequencing etc). One main challenge in systematically implementing those approaches is thus their standardization across observatories. Based on a survey of existing methods and on a 3-year experience in collecting, sequencing and analyzing molecular data, this deliverable is thus dedicated to present the SOPs implemented and tested at NEREA. The SOPs consider a size fractioning of the biological material to avoid biases toward more abundant, smaller organisms such as bacteria. They cover both the highly stable DNA and the less stable RNA and they are essentially an evolution of the ones developed for the highly successful Tara Oceans Expedition and recently updated for the Expedition Mission Microbiomes, an All-Atlantic expedition organised and executed by the EU AtlantECO project. Importantly, they have only slight variations with respect the ones adopted by the network of genomic observatories EMOBON. Discussions are ongoing with EMOBON to perfectly align the protocols. The SOPs are being disseminated via the main national and international networks.

1. Introduction

Marine ecosystems are of vast importance to human and planetary health as they provide a wide range of services [1]. Nevertheless, vast areas of the Earth's oceans have been and are being severely impacted by human activity. Addressing the pressing threats to marine ecosystems requires a deeper and more focused understanding of how multiple stressors affect their ecological diversity and functioning. It also requires enhancing our monitoring capabilities to detect changes and identify near real-time threats to ecosystems and related services, threats such as harmful algal blooms occurring in areas of interest to aquaculture [2].

Recent technological developments in molecular methods such as genomics, transcriptomics and related 'omics' approaches hold great promise in reporting on the status and changes in ubiquitous microbial communities in marine ecosystems, which are intimately linked to ocean biogeochemistry [3]. Indeed, marine omics observation allows the detection of responses to both anthropogenic and natural stressors in great

detail [4], [5]. Not surprisingly, omics approaches are increasingly popular both as additional protocols in oceanographic campaigns and in coastal monitoring. Intensive work is also being done to realise the full potential of omics observations through the global coordination of best practices, FAIR data products and calibration actions with existing ocean observation infrastructures.

To support the full inclusion into routine operations of technological advances in biological oceanography, the G7 working group on the "Future of oceans and seas" in 2016 has proposed an action that, in its essence, consists in augmenting existing marine observatories with novel ecological sensing technologies and know-how. It builds upon existing and developing sensing technologies in biogeochemistry, oceanography, and imaging with the ongoing "-omics revolution" in biology, enabling unprecedented, holistic insight into marine ecosystems.

In particular, omics techniques aim to fill an important gap in existing sensing capabilities: the ability to detect the composition, functions and responses of ecological communities that drive and respond to marine phenomena at different scales and without capacity constraints. Omics data can detect signals across the full range of ecosystem species, enabling multifunctional and cost-effective surveillance of numerous ecosystem properties, including biodiversity, stress, pollution responses and the presence of invasive species [6]. In addition, the data can be used to provide services to society, such as the prediction of hazards such as harmful algal blooms and bioprospecting for the Blue Bioeconomy [7].

To realise their full potential, omics approaches need to be combined and contextualised with high-resolution data from biogeochemical, imaging and physical sensing devices. While this synthesis would allow for a much more holistic view and assessment of marine ecosystems, which is crucial for the future of marine policy, its achievement requires considerable method development work and its testing under operational conditions.

The rationale for upgrading existing observatories is as follows:

(a) Observatories already have operating procedures and regular activities for the collection of data on important environmental parameters on the physical and chemical environment and on the ecological and biological state of the ecosystem (e.g., Chlorophyll, plankton biomasses, etc). It is thus a natural step to augment with advanced biological tools. In turn, the -omics data have a much higher value due to this integration.

b) Building on an existing infrastructure is cost-effective; in addition, pre-existing knowledge of local ecosystems provides a baseline against which enhanced observations can be compared.

b) They are already connected to scientific and monitoring networks and can therefore easily disseminate new advances; moreover, they can make use of existing coordination frameworks to ensure efficiency and avoid duplication.

c) A network of observatories using new technologies, including omics, simultaneously will enable broad support to overcome challenges through accelerated capacity building and knowledge transfer.

d) A network using comparable techniques will enable a large-scale, collective synthesis of ocean conditions, the efficient development of new marine health indicators and, importantly, serve as a site for testing new sensors for robotic monitoring.

Scaling up the activities of an observatory, i.e. transferring the momentum and promise of emerging technology and in particular the omics revolution into the marine monitoring arena, requires going through several incremental steps, such as the creation of coordinating working groups, the development and dissemination of standardised data processing pipelines and sampling protocols, the creation of legal frameworks, the integration of omics data with other marine data, and the development of new indicators. In addition, pilot actions are needed to test and validate methods and technologies.

The task 3.8 of the project was designed with the specific aim of supporting the G7 call for action. The objectives are to: 1) Develop and implement a set of standard operating procedures (SOPs) for long-term omic observation by augmenting well-established marine LTERs; 2) Leverage existing omic coordination networks to disseminate SOPs and solicit international feedback; 3) Contribute to shape international standards to the needs of marine omics observatories; 4) Disseminate network-reviewed SOPs. At this aim a workshop with international experts was organised to align protocol development with existing activities and maximize global impact and synergy. Further, a test of selected omics SOPs (adapted to an LTER setting) was performed through a 12-month implementation. This deliverable is dedicated to the presentation of the resulting SOPs.

In reception of the G7 call for action, the Stazione Zoologica Anton Dohrn (SZN) in February 2019 launched the NEREA (Naples Ecological REsearch for Augmented Observatories; https://www.nerea-observatory.org/) programme, a multidisciplinary, integrative observatory of the water column. It was largely inspired by the success of Tara Oceans, an oceanographic expedition applying state-of-the-art molecular techniques, morphological and environmental data to study the enormous variety of organisms in the marine microbiome [8] (https://fondationtaraocean.org/en/home/). The core idea of NEREA is to create a tight connection among the different components of a modern, holistic research programme on the marine ecosystems. This is translated into a continuous dialogue across disciplines to ensure two-way feedbacks between the in situ, laboratory and in silico exercises.

Importantly, NEREA augments a marine LTER, MareChiara (MC), and builds on the scientific knowledge and technical expertise of this weekly monitoring programme, established in 1984 [9]. Standard MC sampling consists of monitoring the physics and chemistry of the water column at a fixed point in the Gulf of Naples, plus fine taxonomic analysis of phytoplankton and zooplankton communities.

The augmentation mainly consists of adding a size-fractionated sampling of the plankton for the characterization of the community structure and functional state. At this aim, protocols for metabarcoding, metagenomics and metatranscriptomics have been developed, tested and now routinely applied on monthly scales since 2019 (with some pauses due to the pandemics).

NEREA thus produces a large series of datasets from its monthly activities. While the physical and chemical data are mostly processed at SZN, for the -omics data SZN has no capacity for the kind of sequencing required by the actual international standards (Illumina) and has no specific bioinformatic expertise to manage -omics data, especially for eukaryotes. For this reason, and given the long, very successful collaborations (details below) with GENOSCOPE, NEREA has been established with GENOSCOPE as partner since the beginning.

NEREA is currently supported by the EU projects EuroSea and TechOceanS. Both projects provide the financial support for the sequencing of the NEREA metagenomic and metatranscriptomics samples for the period

2019-2022. The activities will continue since three different new EU projects and the Italian PNRR will support the future costs.

Specifically, in EuroSea project NEREA is serving as a reference for “Best practices” for the establishment of a network of -omics enabled, augmented observatories at EU level (Task 3.8). The NEREA -omics SOPs have already been deposited in an open-access online Ocean Best Practices System (OBPS) repository in the OBON collection [10]. The SOPs can be found here: https://repository.oceanbestpractices.org/handle/11329/2375.

In the context of EuroSea, SZN has not only exploited NEREA as a tool for developing best practice but it has also led the scoping workshop that has served to create the EuroGOOS Working group on biological observations (https://eurogoos.eu/biological-observations-working-group/) whose role is to disseminate best practices on molecular and imaging methods in the monitoring of the EU marine ecosystem health.

The NEREA -omics dataset and analyses supported with the EuroSea funds will thus serve to demonstrate the scientific and societal interests in such “augmentation”. Specifically, to demonstrate the added value of the such -omics protocols, SZN will disseminate these best practices at national level (PNRR – National Biodiversity Center - SPOKE 2), at EU level (EuroSea partnership and EuroGOOS WG) and at global level via the participation to the UN Decade of Ocean Sciences (UNDOS) and, specifically, to the global Ocean Biomolecular Observing Network (OBON) framework (NEREA has already obtained the official endorsement by the UNDOS while two coordinators of NEREA are among the founding members of OBON and sit in its OBON steering committee). The synergy with the EU TechOceans project instead aims at introducing new approaches to the monitoring of marine ecosystems (sensors, robots etc). SZN contributes to the development of bioassays for selected toxic species and to host the deployment of new sensors developed by other partners of the projects. The NEREA -omics database and analyses thus served as background information for the sensors' design and deployment. The activity could, in principle, bring to new patented designs.

This Deliverable is dedicated to present the SOPs derived from the NEREA experience. As mentioned, given the recent advancements in fundamental biology due to the introduction of genomic approaches several research institutions are moving into introducing new -omics protocols to the regular sampling of marine ecosystems and specifically of the oceanic microbiomes (viruses, bacteria, phyto) and to their main predators (zooplankton). In practice this corresponds to filter the seawater using different tools for the water collection (e.g., Niskins bottles, pumping systems, nets) and complementary filter sizes. The filters are then stored in freezers and subsequently sent to the sequencing center. There, protocols are applied to extract DNA (for taxonomy and to identify potential functions) as well as the RNA (to monitor for the genes activity in various abiotic and biotic conditions) and then to proceed to the sequencing. The resulting raw sequences (fragments of genomes) are then processed using bioinformatic tools to produce datasets for direct scientific use. Essentially, three main data sets are produced: 1) a taxonomic data set based on metabarcoding (sequencing of selected genome regions); 2) a metagenomic dataset, made of “putative” genes (with, when possible, a gene-by-gene functional and taxonomic assignations); 3) a RNA-based metatrascriptomics dataset, made of “putative” functional genes expressions.

There are still several limitations. For instance, the scarcity of known genomes strongly limits the taxonomic and functional assignments, which sometimes do cover only 50% of the datasets. This is one of the reasons why it is cost-effective augmenting observatories (where significant knowledge is present, including genomes of locally relevant species). In addition, while the standardization of protocols and processing methods is

crucial for data comparability, it is far from being achieved. This is particularly relevant given the intrinsic compositional nature of the data.

The main issue with the standardization is in the choice of the size fractioning of the organisms. In the oceans, the abundance of various kinds of organisms varies by several orders of magnitude, with bacteria and nano-eukaryotes being far more abundant than large phytoplankton and zooplankton. This means that by using a single filter size, generally very small, the larger organisms represent only a small (if not tiny) fraction of the abundances. Correspondingly, in most cases, the sequencing produces data that are statistically reliable only for the smaller organisms. A good coverage of the entire microscopic food web is thus obtained only by size fractioning the organisms and then proceeding to separated sequencing exercises. While being obviously better, this approach has two important drawbacks: 1) it requires a much more complicated set of protocols to be executed on board (the filtration has to be performed as soon as possible after the sampling at sea), with Niskins or pumps for collecting small organisms and the use of nets for the larger ones; 2) it clearly increases the cost of the sequencing. It is thus strategic choice, which has huge implications (It is almost impossible to merge properly data collected with different size criteria). In the case of the augmentation of the observatory, the choice was made of using the full size fractioning (to fully cover the ecosystem) while reducing the frequency (from weekly to monthly) to guarantee the financial sustainability of the sampling over time. To maximise the production on knowledge on the system's dynamics, seasonally additional stations are performed over small transects across environmental gradients (river-impacted waters vs open waters, mostly).

There are many other challenges related to these kinds of data, challenges such as defying the optimal amount of water volume to be filtered and to identify the best storage process. In addition, the association of the omics-derived data with estimates of biomass, biovolumes and carbon contents remains an open challenge. Another challenge consists in the complexity of the data produced and the need to match the standardized oceanographic data categories since, among others, the biological community is not tightly linked to the (operational) oceanographic one. At the aim of easing a dialog, a dedicated workshop will be organized in the following months. Moreover, the establishment of the EuroGOOS WG on Biological observations is certainly a positive step in the direction of filling these gaps, gaps which can be additionally reduced by exploiting the EuroGOOS mechanisms such as the Integration Workshops.

Finally, it is important to mention that while the EuroSea activities were on going, EMBRC launched in summer 2021 a new project, the European Marine Omics Biodiversity Observation Network (EMO BON), with the aim of enhancing the European contribution to global genomic observation efforts. The network is made of the EMBRC marine stations (which include SZN). In each marine station a monthly sampling exercise is performed using dedicated protocols which cover also the plankton. Born by having NEREA as one source of inspiration and in collaboration with SZN, the network (https://www.embrc.eu/services/emo-bon) will ensure steady, continuous generation of 'baseline' data on biodiversity at EMBRC sites following FAIR (Findable, Accessible, Interoperable, and Reusable) data principles. A dedicated committee, composed also by one of the NEREA coordinators, has designed the SOPs, which are very closely related to NEREA SOPs in being both derived by the experience of Tara Oceans. It is to note that NEREA is the only site that has already collected and sequenced a full one year of full -omics data. The EMOBON consortium is starting now the sequencing and, for a while, it is going to focus on metabarcoding and metagenomics only.

About at the same time, a series of large-scale, genomic-enabled expeditions were organised by an EuroSea EU sister project (AtlantECO). This effort was supported by the production of SOPs, which are also an evolution of the Tara Oceans' one.

2. Standard operating Procedure for OMICS size fractionation

Here the details of the SOPs for augmenting the monitoring at a coastal observatory are presented. The oceanographic conditions at the LTER sampling site in the Gulf of Naples are characterized by a significant seasonal cycle in the atmospheric and land forcings (river input) and by the continuous reversing of the Gulf inner horizontal circulation [11,12]. As a result, an important seasonal cycle characterizes the plankton community, with significant spring blooms, dominated by large diatoms, and with oligotrophic summers (the background state of the offshore waters is typically sub-tropical) with a significant accumulation of the biomass in the zooplankton compartment (copepods, mostly) [13,14,15]. Small flagellates dominate in winter, when light is limiting and/or vertical mixing is very active.

In order to fully cover the plankton size spectra, the main strategic choice was to separate the organisms on the basis of their size. Specifically, following common practices, the choice was to split the samples into these size fractions: 0.2-3, 3-20, 20-200 microns. A larger size fraction was also considered, with protocols which are essentially the same as EMOBON.

The full list of the protocols can be found here: https://www.nerea-observatory.org/nereasampling.

The protocols are very close to the EMOBON protocols for plankton -omics. The latter can be found here, with details on the implementation of the SOPs: https://www.embrc.eu/newsroom/publications/european-marine-omics-biodiversity-observation-network-emo-bon-handbook.

For completeness, the EU AtlantECO protocols, which are the result of a global consultation, can be found here: https://zenodo.org/record/6956974. It is to note that the latter ones were developed for oceanic expeditions.

Here below the details for the execution of the SOPs.

2.1. Sample collection

- Sample seawater from the selected depth using Niskin bottles previously thoroughly rinsed with tap water;
- A total of 10 L seawater are collected from the Niskin bottle after sieving through 200 and 20 µm mesh size nylon sieves, sequentially;
- The 200 µm sieve (Fig. 1) is placed on top of a 20 µm net inside a funnel (Fig. 2);
- The sieved seawater is collected into an acid-cleaned graduated tank with a tap (named IN-tank);
- Make sure that the tap of the IN-tank is closed and turned up, before collecting the water;
- After sieving 10 L seawater, place the 20 µm sieve (from the funnel) in a beaker with some filtered seawater to keep it humid;
- Carry the IN-tank to the bench to start filtration.

2.2. Sample filtration: Size fractions 0.2-3 and 3-20 micron

The filtration apparatus is composed by two 142 mm filter holders (tripods), connected with appropriate silicon tubes, an EZ-Stream Merck Millipore vacuum pump, and a collecting tank, named OUT-tank, with a tap (Figure 1). The activities have to be performed one after the other. It is particularly important to operate the filtration for the metaT samples within hours from the collection.

Filters are Millipore Isopore TSTP 14250 (3 µm) and Millipore Isopore GTTP14250 (0.22 µm).

Warnings:

- It is crucial that the tank with ingoing water (called IN-tank) is placed higher than the first tripod which, in turn, is placed higher than the second tripod;
- It is important that the pump is placed at a safe distance from the tripods, to avoid contact between water and sockets;
- Correct placement of the pump is shown in Figure 1, where the pump is located on a support (in this case a chair), and behind the workbench;
- The pump is then connected through a tube to the tank of the outgoing water (OUT-tank), placed on the floor.

Figure 1. Pump placement in the wet lab on board

Filtration setup

a) Always wear gloves;
b) Place clean paper on the benchtop;
c) Clean tweezers with ethanol;
d) Open the tripods on the benchtop, removing the top and placing it upside to avoid contact with surfaces;
e) Place the 3 µm membrane filter in the higher tripod and the 0.22 µm filter in the lower one. Keep the same orientation as they are in the box and discard the blue paper inter-filters;
f) Filters are to be placed on the blue plastic support of the tripod;
g) Once filters are placed inside each tripod, close them, tighten the hand knobs gradually and simultaneously (this facilitates the correct sealing of the tripod, avoiding spillovers);
h) Connect the inlet and outlet tubes;
i) Place the tripods as indicated in point I and shown in Fig.1 and set up the connections among the different components of the filtration apparatus.

Filtration (reduce the filtration time to a max of 30')

a) Before starting filtration, make sure to mark the cryovials appropriately;
b) MC is the code for MareChiara, S for Sarno river, C for Capri stations;
c) The code for the fraction 0.22-3 µm (Free Bacteria) is "NRFB+CODE OF STATION+SEQUENTIAL NUMBER OF SAMPLING AT THE SPECIFIC STATION SITE+SAMPLE NUMBER" (e.g. NRFB C5 #82);
d) The code for the fraction 3-20 µm (NanoEukaryotes) is "NRNE+ CODE OF STATION+SEQUENTIAL NUMBER" (e.g. NRNE C5 #82);
e) Connect the IN-tank to the higher tripod with a tube and open the tap, allowing the complete filling of the first connecting tube;
f) Make sure that the pump is connected with a tube to the OUT-tank;
g) See the scheme below:
h) Turn the pump on;
i) Always initiate the filtration with the filter holder vents open. Wait until air bubbles are expulsed and close the vents;
j) Filter 5 L, then close the tap of the IN-tank and let all the remaining water to be filtered through;
k) Once no more water is left in the tubes, turn the pump off;
l) Disconnect the tripods and put them on the bench;
m) Open the higher tripod, remove the top and position it upside down on the bench, avoiding the contact with surfaces;
n) Use the tweezers to handle the filter;
o) Fold the filter in half repeatedly by folding in a cylinder to fit in a vial;
p) Carefully place the filter in a 5 mL cryovial and gently push it to the bottom;
q) Close the cryovial and quickly place it in the liquid nitrogen container;
r) This is the ii fraction (3-20 µm);
s) The code for the fraction 3-20 µm (NanoEukaryotes) is "NRNE+ CODE OF STATION+SEQUENTIAL NUMBER" (e.g. NRNE C5 #82);
t) Repeat points i to m for the lower tripod;
u) This is the i fraction (0.2 – 3 µm);
v) The code for the fraction 0.22-3 µm (Free Bacteria) is "NRFB+CODE OF STATION+SEQUENTIAL NUMBER OF SAMPLING AT THE SPECIFIC STATION SITE+SAMPLE NUMBER" (e.g. NRFB C5 #82);
w) Repeat points a) to r) for the remaining 5 L (second replicate);
x) Store the OUT-tank in the shade; the 0.22 µm filtered seawater will be used for virus-DNA (link to the protocol).

2.3. Sample filtration: Size fraction 20-200 micron

This is done after completing the process for 0.2-3 and 3-20 micron fractions.

a) Always wear gloves;
b) Clean the workbench and replace the paper;
c) Wash the IN-tank and the tripods with DW;
d) Use only one tripod
e) Clean tweezers with ethanol;
f) Place a 3 µm filter as described in A-1;

g) Place the tripod on the stool;
h) Take the beaker with the 20 µm sieve from the initial collection (Water sampling, f);
i) Rinse the 20 µm nylon sieve using the filtered water from the OUT-tank;
j) Pour the remaining filtered water from the OUT-tank into the IN-tank, together with the water used to rinse the nylon sieve in h);
k) Connect the tripod to the pump and to the tubes;
l) Make sure that the pump is connected with a tube to the OUT-tank;
m) Turn the pump on;
n) Always initiate the filtration with the filter holder vents open. Wait until air bubbles are expulsed and close the vents;
o) Filter 5 L, then close the tap of the IN-tank and let all the remaining water to be filtered through;
p) Once no more water is left in the tubes, turn the pump off;
q) Disconnect the tripod and put it on the bench;
r) Open the tripod, remove the top and position it upside down on the bench, avoiding the contact with surfaces;
s) Use the clean tweezers to handle the filter;
t) Fold the filter in half repeatedly by folding in a cylinder;
u) Carefully place the filter in a 5 mL cryovial and gently push it to the bottom;
v) Close the cryovial and quickly place it in the liquid nitrogen container;
w) This is the fraction 3-20 micron;
x) THE CODE FOR THIS FRACTION IS NRME+CODE OF STATION+SEQUENTIAL NUMBER;
y) Repeat points I to u for the remaining 5L (second replicate);

Once in the lab, the cryovials are stored at -80 °C.

The OUT-tank, containing filtered seawater is used to collect samples for viral diversity (see appropriate SOP).

3. List of materials, equipment and supplies

- Filtration Apparatus (2 Tripods + Pump + Power Supply)
- Stool And Wood Base
- Silicon Tubes With Different Sections For Pump-Filtration Apparatus- Tank
- 10 L Graduated Tank 10 With Tap (In-Tank)
- 10 L Tank With A Tap (Out-Tank)
- 20 µm Mesh Size Nylon Sieve
- Plastic Funnel
- 200 µm Mesh Size Sieve Mounted On A Plastic Container (Fig. 1)
- 2 Tweezers
- Polycarbonate Filters 142 Mm Diameter 0.22 Millipore Isopore Gttp14250) and 3 µm (Millipore Isopore Tstp 14250) Pore Size
- Distilled Water
- Ethanol
- 2l Plastic Beakers

- Gloves
- Table Paper
- Six 5 Ml Cryovials (Duplicate Vials For Each Of The Three Size Fraction)
- Marker Pen Resistant To Liquid Nitrogen (Sharpie Or Similar)
- Liquid Nitrogen

4. Sequencing

To validate the SOPs a pilot operational phase was designed and executed in 2019 as part of the activites of the observatory. Monthly samples collected in 2019 were subsequently sequenced at Genoscope. While the pandemic has significantly delayed the data production and the application of the bioinformatic pipelines, it can be reported here that except for two cases (the amount of living material on the filter was too scarce), all the DNA and RNA extractions was successful, confirming the correctness of the choice of the filtered volumes.

A total of 58 metagenomic samples collected from April 2019 to January 2020 and representing different sites, size fractions and depths, were sequenced and analyzed to build metagenome-assembled genomes (MAGs). 33 out of the 58 metagenomic samples had the corresponding metatranscriptomic data. Metatranscriptomic reads have been assembled to build a catalogue of NEREA expressed genes.

Preliminary analyses of the taxonomic content show a good coverage of the local trophic web. Additionally, the metagenomic data are rich enough to allow mapping them onto the genomes of selected model species, even when they were not abundant in the water column.

5. Conclusion

Life underpins marine ecosystems, but our ability to observe it, monitor it and understand its functioning is still very limited. The enhancement of marine observatories with new technologies such as multi-omics is already improving the observation of life forms, from bacteria to tuna, allowing unprecedented insight into the structure and function of marine systems [16,17]. Much remains to be done, however, to enable the inclusion of these methods in a systematic, efficient and sustainable manner.

To this aim, observatories are currently being augmented by adding, among other technological novelties, -omics protocols for characterising and monitoring the plankton communities, in order to assess the ever-changing diversity and the associated metabolic states. Given the many challenges associated with the use of such methods, the augmentation of selected observatories (having a strong background knowledge of the systems and, possibly, also facilities for laboratory and in silico ancillary studies) is a prerequisite for the setting up and wide adoption of standardised SOPs. These recent technological developments in molecular methods such as genomics, transcriptomics and related ‘omics’ approaches are becoming more popular. Thus, the establishment of SOPs at these very first steps of this global tendency is very important in order to avoid discrepancies, and maximise compatibility and in general the exploitation of the data.

Building upon the systems know-how and technical expertise of an almost 30-year old LTER [9], the NEREA augmented observatory has been used as a context to test and refine SOPs of the augmentation with -omics of plankton regular monitoring. The exercise has been conducted while keeping a tight coordination with international initiatives such as the recently created EMOBON-EMBRC network of genomic observatories and the EU sister project AtlantECO, dedicated to the assessment of the Atlantic Ocean microbiome health status and functioning.

Much work remains to be done. It is fundamental to create a database of the genomes of the species of ecological and (potentially) economic interest. A proper pipeline that connects the production of new data to the existing oceanographic databases has also to be established. While being well established, the implementation of the SOPs here proposed imply a significant amount of human and financial resources. This may limit their application at global scales and, thus, it is indeed necessary to proceed with further developments to have versions of the SOPs that are more broadly of interest while being more informative than the classical low profiles ones (generally limited to metabarcoding).

Regarding the dissemination of the SOPs within coordination actions, OBON is the first choice. Further, dissemination is occurring via meetings, working groups (mostly the EuroGOOS WG) and via the preparation of scientific publications.

The existence of EMOBON has created an important reference at EU and at global scales. EMOBON has also addressed issues such as the ABS, biobanking and the legal aspects related to the sampling of national genetic resources. Nevertheless, it is a consortium made of the EMBRC marine stations and the inclusion of new observatories, at the moment, can only pass via the adhesion to the ESFRI infrastructure, a complex process indeed.

Finally, these basic -omics exercises were first introduced in oceanography more than 10 years ago [18]. Meanwhile, many more approaches in advanced biology are being transferred to operational oceanography. Regarding the -omics, it has to be reminded that the analysis of the DNA and even of the RNA (the actual gene expressions) provides only hints about the “potential” status of the ecosystems since there is not a clear, univocal correspondence between this genomic information and the actual production of metabolites. This is why metaproteomics and metabolomics, while in their infancy as operational tools, are very promising to fully inspect the actual metabolic status of the ecosystems [19]. Here again the augmented observatories (including NEREA) are playing an important role for the development of suitable SOPs. Further, the routine inclusion of single cell -omics as well as the derivation of the species-specific genomes from metagenomic data are paving the way to link evolution with ecology. Given the fast reproduction rates of most planktonic forms of life, these approaches will allow inspecting eco-evolutionary responses to, e.g., climate changes. It is thus very timely that the international community is organising a new decadal, global programme, Biogeoscapes (https://biogeoscapes.org/), which will deeply contribute to an international standardization of SOPs as well as to fully integrating advanced biological approaches into oceanography at all levels, from the marine stations to the large-scale cruises.

6. References

1 Barbier, E. B. Marine ecosystem services. *Curr. Biol.* **27**, R507-R510 (2017).

2 Cavicchioli, R. *et al.* Scientists' warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.*, doi:10.1038/s41579-019-0222-5 (2019).

3 Trevathan-Tackett, S. M. *et al.* A horizon scan of priorities for coastal marine microbiome research. *Nature Ecology & Evolution* **3**, 1509-1520, doi:10.1038/s41559-019-0999-7 (2019).

4 Pawlowski, J. *et al.* The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total Environ.* **637-638**, 1295-1310, doi:10.1016/j.scitotenv.2018.05.002 (2018).

5 Mock, T. *et al.* Bridging the gap between omics and earth system science to better understand how environmental change impacts marine microbes. *Global Change Biol.* **22**, 61-75, doi:10.1111/gcb.12983 (2016).

6 Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, doi: 10.1126/science.1261359, doi:10.1126/science.1261359 (2015).

7 Jeffery, N. W. *et al.* Application of Omics Tools in Designing and Monitoring Marine Protected Areas For a Sustainable Blue Economy. *Frontiers in Genetics* **13**, doi:10.3389/fgene.2022.886494 (2022).

8 Karsenti, E. *et al.* A Holistic Approach to Marine Eco-Systems Biology. *PLoS Biology* **9**, doi:e100117710.1371/journal.pbio.1001177 (2011).

9 Zingone, A. *et al.* Time series and beyond: multifaceted plankton research at a marine Mediterranean LTER site. *Nature Conservation* **34**, 273, doi:10.3897/natureconservation.34.30789 (2019).

10 Trano, A.C. Collection standard operating procedures: Omics size fractionation. Version 1.00. Naples, Italy, Naples Ecological Research Augmented Observatory (NEREA), 5pp. doi:10.25607/OBP-1963 (2021).

11 Kokoszka, F., Le Roux, B., Iudicone, D., Conversano, F. & d'Alcala, M. R. Long-term variability of the coastal ocean stratification in the Gulf of Naples: Two decades of monitoring the marine ecosystem at the LTER-MC site, between land and open Mediterranean Sea. *Mar. Ecol.-Evol. Persp.* **44**, doi:10.1111/maec.12725 (2023).

12 Romillac, N. *et al.* Interplay among anthropogenic impact, climate change, and internal dynamics in driving nutrient and phytoplankton biomass in the Gulf of Naples. *Mar. Ecol.-Evol. Persp.* **44**, doi:10.1111/maec.12754 (2023).

13 Longobardi, L., Dubroca, L., Margiotta, F., Sarno, D. & Zingone, A. Photoperiod-driven rhythms reveal multi-decadal stability of phytoplankton communities in a highly fluctuating coastal environment. *Sci. Rep.* **12**, 1-15 (2022).

14 Zingone, A., Tortora, C., D'Alelio, D., Margiotta, F. & Sarno, D. Assembly rules vary seasonally in stable phytoplankton associations of the Gulf of Naples (Mediterranean Sea). *Mar. Ecol.-Evol. Persp.* **44**, doi:10.1111/maec.12730 (2023).

15 Mazzocchi, M. G. *et al.* Coastal mesozooplankton respond to decadal environmental changes via community restructuring. *Mar. Ecol.-Evol. Persp.* **44**, doi:10.1111/maec.12746 (2023).

16 Davies, N. *et al.* The founding charter of the Genomic Observatories Network. *GigaScience* **3** (2014).

17 Santi, I. *et al.* European marine omics biodiversity observation network: a strategic outline for the implementation of omics approaches in ocean observation. *Frontiers in Marine Science* **10**, doi:10.3389/fmars.2023.1118120 (2023).

18 Bowler, C., Karl, D. M. & Colwell, R. R. Microbial oceanography in a sea of opportunity. *Nature* **459**, 180-184 (2009).

19 Schorn, M. A. *et al.* A community resource for paired genomic and metabolomic data mining. *Nat. Chem. Biol.* **17**, 363-368, doi:10.1038/s41589-020-00724-z (2021).