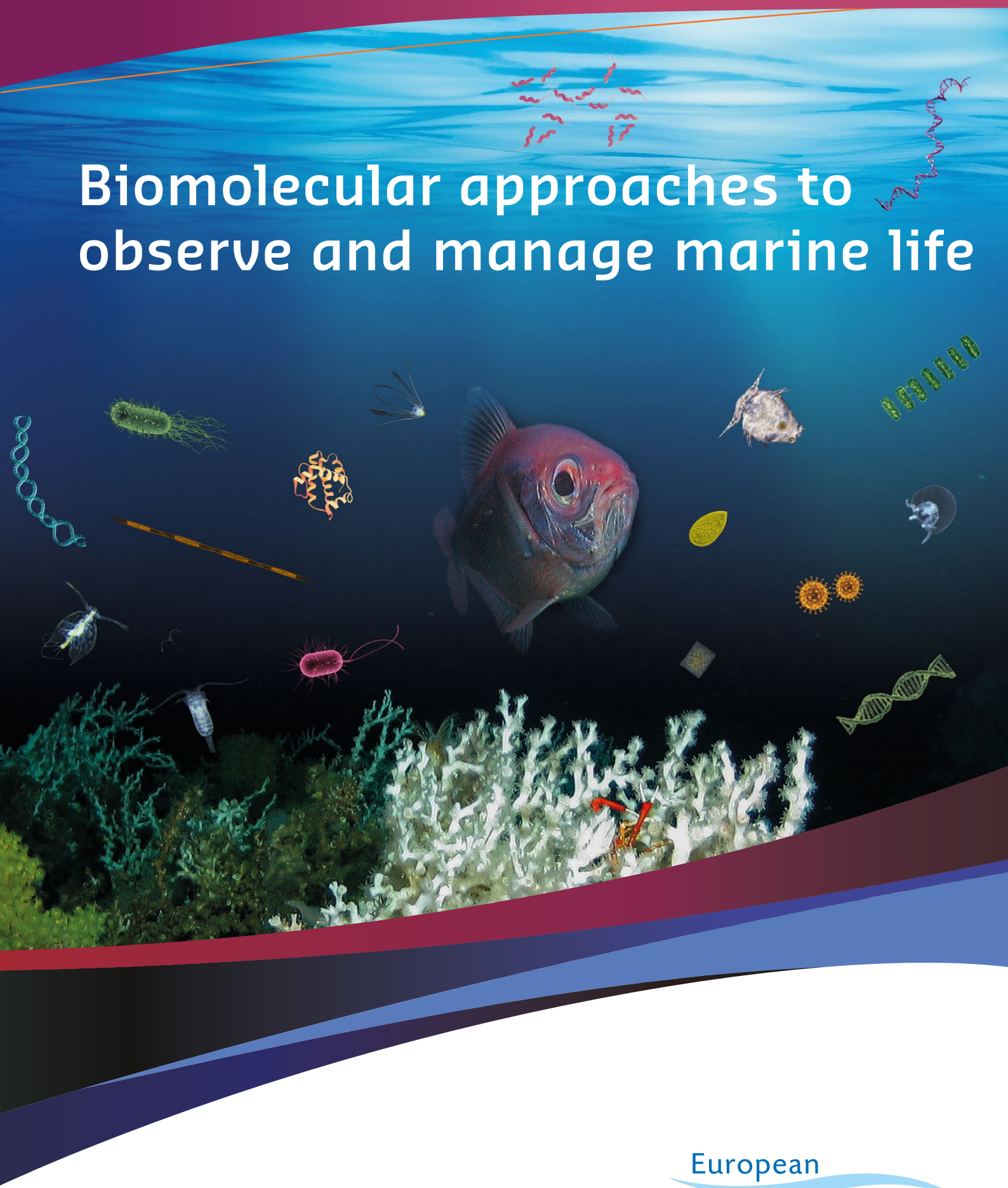
Future Science Brief Nº 14 November 2026

Biomolecular approaches to observe and manage marine life



www.marineboard.eu

European MARINE BOARD
Advancing Seas & Ocean Science

European Marine Board IVZW

The European Marine Board (EMB) provides a pan-European platform for its Member organisations to develop common priorities, advance marine research, and to bridge the gap between science and policy in order to meet future marine science challenges and opportunities.

The EMB is an independent and self-sustaining science-policy interface organisation that currently represents 37 Member organisations from 19 European countries. It was established in 1995 to facilitate enhanced cooperation between European marine science organisations towards the development of a common vision on the strategic research priorities for marine science in Europe. The EMB promotes and supports knowledge transfer for improved leadership in European marine research. Its membership includes major national marine or oceanographic institutes, research funding agencies and national consortia of universities with a strong marine research focus. Adopting a strategic role, the EMB serves its Member organisations by providing a forum within which marine research policy advice is developed and conveyed to national agencies and to the European Commission, with the objective of promoting the need for, and quality of, European marine research.

www.marineboard.eu

European Marine Board Member Organisations

European Marine Board IVZW Future Science Brief N° 14

This Future Science Brief is a result of the work of the European Marine Board Expert Working Group on Biomolecular Approaches to the Observation and Management of Marine Life. See Annex 1 for the list and affiliations of the Working Group Members.

Working Group Chairs
Louise Allcock, Paloma Morán

Contributing Authors
Ana Baričević, Chris Bowler, Damien Eveillard, Pascal Hablützel, Bruno Louro, Antonella Penna, Benjamin Pontiller, Julie Robidart, Jorge Juan Santos Bruña, Elena Stoica, Agnes Weiner

Series Editor
Sheila J. J. Heymans

Publication Editors
Britt Alexander, Zoé Levit, Ana Rodriguez Perez, Paula Kellett, Ángel Muñiz Piniella, Sheila J. J. Heymans

External Reviewers
Angel Borja, Kelly Goodwin

Internal review process
The content of this document has been subject to internal review and approval by the European Marine Board Member organisations.

Suggested reference
Allcock, A.L., Morán, P., Baričević, A., Bowler, C, Eveillard, D., Hablützel, P., Louro, B., Penna, A., Pontiller, B., Robidart, J., Santos Bruña, J. J., Stoica, E. Weiner, A. (2026) Biomolecular Approaches to Observe and Manage Marine Life. Alexander, B., Levit, Z., Rodriguez Perez, A., Kellett, P., Muñiz Piniella, A., Heymans, J. J. [Eds.] Future Science Brief N°. 14 of the European Marine Board, Ostend, Belgium. ISSN: 2593-5232, ISBN: 9789464206432, DOI: 10.5281/zenodo.23419880

www.marineboard.eu
info@marineboard.eu

Design
Zoeck

Front cover
Composition by Zoé Levit. Image Credits: MARUM, Center for Marine Environmental Sciences, University of Bremen (CC BY 4.0); © kristobal22, iNaturalist (CC BY NC 4.0); Pixabay; © ARTIS, Freek Wagenaar; Ruđer Bošković Institute, Center for Marine Research

First edition, November 2026

This document in licensed under CC BY 4.0.

In this document, Artificial Intelligence-powered tools have been used to assist with language editing and the design of some figures (see Annex 2 tools used and for which purposes). AI has not been used for any key writing task, such as producing scientific insights, creating a literature review, drawing scientific conclusions or providing recommendations.

Foreword



The Ocean is undergoing rapid and widespread change and there is a critical need for comprehensive biodiversity assessments to support effective marine ecosystem-based management. The protection and sustainable use of biodiversity is the centre of global and European Ocean governance agendas. The Agreement on Marine Biological Diversity of Areas Beyond National Jurisdiction recognises the need for advances in environmental assessment tools, such as biomolecular methods, for informed decisions on Marine Protected Areas, Environmental Impact Assessments, and to establish best practices for Access to Marine Genetic Resources. In addition, several European directives and strategies set legal requirements for comprehensive biodiversity and ecosystem monitoring, including the Marine Strategy Framework Directive and the 2030 Biodiversity Strategy. There is also increasing demand for improved observations to support these goals, with the Global Ocean Observing System advocating for the wider use of biomolecular technologies.

Technological advances over the last 20 years mean that biomolecular tools are now more accessible and cost-effective, and their use in marine research has allowed high-resolution insight into ecosystems and biodiversity. However, there is now a need to scale-up the use of these tools to complement traditional methods in conventional monitoring programmes, and to feed data into Digital Twins of the Ocean and decision-making frameworks. The European Commission's OceanEye initiative will be critical to support the innovation required for biomolecular tools to become an integrated component of Ocean observation systems.

At the EMB Autumn Plenary meeting in 2024, the topic of "Biomolecular approaches to the observation and management of marine life" was selected as a new Working Group, and the kick-off meeting was held in October 2025 at the Ryan Institute, University of Galway, Ireland, hosted by Prof. Louise Allcock, the Working Group Chair. This Future Science Brief examines state-of-the-art biomolecular tools and how they can transform our ability to detect and monitor marine life. It presents recommendations for the scientific advances needed for their wider uptake in both research and policy, and for accurate data interpretation, methodological standardisation, advanced data infrastructure, and capacity development. As we move towards the latter part of this decade, this Future Science Brief is an important contribution to the UN Decade of Ocean Science for Sustainable Development and the EU's Mission to Restore Our Ocean and Waters by 2030.

On behalf of EMB, I would like to thank the Working Group members for their dedication and enthusiasm in writing this document. In particular, I would like to thank Louise Allcock and Paloma Morán for their leadership, the external reviewers for their insights and perspectives and internal reviewers from EMB member organisations for their constructive feedback. Finally, would like to thank the EMB Secretariat, in particular Britt Alexander, for her immense commitment to the coordination of the working group, publication editing and onward communication and dissemination.

As this is the last EMB publication that will be launched under the leadership of the current Executive Director, Prof. Sheila Heymans, I wish to take the opportunity to thank her for her unwavering contributions to the EMB documents and foresight exercises over many years. We hope she will remain closely connected to the work which she has undertaken with the Secretariat and she leaves a great legacy and platform for continued impact in the European marine landscape. From all of the EMB members, thank you Sheila!

Fiona Grant
Chair, European Marine Board
November 2026

Table of Contents

Executive Summary 7

1. Introduction 10
   - 1.1. Traditional observation methods 10
   - 1.2. The biomolecular approach 12
   - 1.3. Application of the biomolecular approach to policy 13
2. Applications for research, management and conservation 14
   - 2.1. Summary 14
   - 2.2. “Who’s there?” From presence to abundance in marine biodiversity assessments 14
   - 2.3. “Who does what?” From presence of organisms to marine molecular function assessments 18
   - 2.4. Investigating the Ocean across scales 20
   - 2.5. Monitoring the Ocean and predicting its future 23
     - 2.5.1. Molecular technologies for the sustainable use of our Ocean 23
     - 2.5.2. From reactive surveillance to predictive pathogen tracking 23
     - 2.5.3. Anticipating critical transitions and tipping points 23
     - 2.5.4. Assessing acclimation and adaptation under environmental change 24
     - 2.5.5. Redefining ecosystem services through functional genomics 24
   - 2.6. From molecular metrics to policy implementation 25
   - 2.7. Recommendations 28
3. State of the art in biomolecular methods 29
   - 3.1. Summary 29
   - 3.2. Sources of molecular material 29
   - 3.3. Methodological state of the art 29
     - 3.3.1. PCR-based methods 34
       - 3.3.1.1. Quantitative PCR 34
       - 3.3.1.2. Digital PCR and droplet digital PCR 35
       - 3.3.1.3. RPA-CRISPR 35
       - 3.3.1.4. (Meta)barcoding 36
     - 3.3.2. -Omics approaches 37
       - 3.3.2.1. (Meta)genomics 37
       - 3.3.2.2. (Meta)transcriptomics 38
       - 3.3.2.3. (Meta)proteomics 39
       - 3.3.2.4. Epigenetics / Epigenomics 40
       - 3.3.2.5. Metabolomics 41
     - 3.3.3. Reference databases 42
     - 3.3.4. Bioinformatics and modelling for generating knowledge from data 43
   - 3.4. Recommendations 44

4. Emerging technologies, and research and development needs 45
4.1. Summary 45
4.2. Deployable autonomous technologies 45
4.2.1. Autosamplers and passive samplers 45
4.2.2. *In situ* sensors 47
4.2.3. Adaptive sampling on autonomous platforms 48
4.2.4. Technological development for deployable technologies: A future outlook 48
4.3. Emerging molecular laboratory technologies: Long-read and miniaturised sequencing 49
4.4. Multi-omics and scaling augmented observatories 50
4.5. Computational methodologies to enrich data interpretation 50
4.6. A vision for molecular biological applications for the Ocean 51
4.7. Recommendations 53
5. Operationalisation, standardisation and FAIR data 54
5.1. Summary 54
5.2. Operationalisation 54
5.3. Standardisation 54
5.4. FAIR and open data 57
5.5. Data processing and workflow documentation 57
5.6. The aggregation paradox: Quality over quantity 59
5.7. Data products and multi-domain integration 59
5.8. Recommendations 60
6. Capacity building and Ocean literacy 61
6.1. Summary 61
6.2. Capacity building 61
6.3. Ocean literacy 62
6.4. Recommendations 63
7. Overarching recommendations 64
List of abbreviations and acronyms 69
Glossary 73
References 79
Annex 1 89
Annex 2 89

Executive summary

Biomolecular approaches, including -omics and eDNA sampling, are becoming increasingly important components of Ocean observation systems. When combined with traditional approaches, "augmented observatories" have potential to provide high-resolution biodiversity and ecosystem insights for the predictive and mechanistic understanding of the Ocean that is required for conservation and ecosystem-based management under rapid environmental change. Scalable and reproducible biomolecular metrics can help establish reliable biological baselines; improve the scientific basis for monitoring protected species; provide early warning for pathogens, invasive species and harmful algal blooms; assess environmental impacts of human activities and climate change; and support conservation measures such as Marine Protected Areas. Biomolecular methods need to be scaled up from research applications to routine monitoring to improve efficiency, accuracy and cost-effectiveness, thereby transforming the implementation of European Directives and international agreements on biodiversity.

This document presents current and emerging applications of biomolecular tools in research and policy contexts (Chapter 2), and provides an overview of state-of-the-art methods, their quality considerations, costs, and comparative advantages and limitations (Chapter 3). Emerging field and laboratory technologies are discussed (Chapter 4), including autonomous samplers, *in situ* sensors, miniaturised platforms, multi-omics, and computational approaches. The need for operationalisation and standardisation in biomolecular methods and data is presented (Chapter 5) and interventions to increase capacity building and Ocean literacy to ensure implementation of biomolecular approaches at scale are discussed (Chapter 6). Each chapter of the document ends with specific recommendations for research funders, infrastructures, policymakers and scientists, which are summarised below.

Recommendations for research funders:

(1) **Fund co-designed research, including the exploitation of large, interdisciplinary datasets** that link biomolecular methods to policy applications. This includes research on molecular traits, -omics functional knowledge and plankton metabolism. Machine-readable indicators will enable data integration into climate models and Digital Twins of the Ocean to improve future spatial climate scenario projections;

(2) **Fund exploratory and curiosity driven fundamental science to investigate baseline molecular mechanisms** in marine organisms and Ocean ecosystems to advance understanding of ecosystem function and dynamics. Research is needed to: characterise genes with unknown functions; determine the speed of DNA degradation; analyse existing datasets; advance multi-omics; better link organisms, genes and functions; and improve ecological models;

(3) **Invest in accessible and interoperable computational and digital infrastructure**, including repositories, visualisation tools and federated reference libraries. Support is needed for data generation, storage, curation, quality control and adherence to FAIR metadata standards, while enabling long-term monitoring and compliance with the Nagoya Protocol and the BBNJ Agreement. This will facilitate integration of biomolecular data into Digital Twins of the Ocean for EU policy implementation;

(4) **Fund the targeted expansion and improvement of molecular reference databases through sequencing campaigns** that generate barcodes, genomes, transcriptomes and proteomes from organisms that are currently not, or poorly, represented in reference databases;

(5) **Support long-term -omics observatories alongside traditional sampling and monitoring networks (i.e. augmented observatories)** that feed into holistic ecosystem monitoring and assessments;

(6) **Fund innovation in automated and miniaturised *in situ* biomolecular samplers and sensors,** including scaling-up and commercialisation of existing technologies, and the expansion of AI for rapid sequence identification and association of -omics data with other environmental parameters to predict planetary-scale trends;

(7) **Enforce rigorous standardisation for biomolecular protocols and ethical open-access frameworks** across the entire -omics workflow to ensure data reuse and interoperability. Embed FAIR and CARE principles in research funding;

(8) **Fund interdisciplinary biomolecular standardisation task forces** within international research networks to ensure that biomolecular discoveries translate into actionable management knowledge; and

9) **Fund sustained training, collaboration and stakeholder engagement** connecting technology developers, data experts and practitioners to improve data and metadata quality, develop AI-ready datasets, and support long-term citizen science. Provide sustained support and resources to developing regions to ensure expertise advances equitably.

Recommendations for policymakers:

(1) **Integrate insights on biodiversity and ecosystem function from biomolecular data into the design of protected areas and restoration initiatives** to support EU policies including the Habitats Directive and Nature Restoration Regulation;

(2) **Embrace novel frameworks to value ecosystem services based on molecular functions;**

(3) **Commission solutions for emergent environmental issues from data-driven approaches** including big data analyses and large-scale models, and adopt management practices that are backed by large and interdisciplinary datasets; and

(4) **Promote the adoption of FAIR and CARE principles for data collected within monitoring programmes** and the publication of data as open as possible and as restricted as necessary.



Contribution to the EU Mission: Restore our Ocean and Waters

This Future Science Brief and its recommendations support the objectives of the EU Mission: Restore our Ocean and Waters in the following ways:

- **"Protect and restore marine and freshwater ecosystems and biodiversity"** by recommending biomolecular methods for enhancing species inventories through the detection of rare, cryptic and invasive species, and for monitoring changes in biodiversity in response to human activities, climate change, and restoration.
- **"Prevent and eliminate pollution of our Ocean, seas and waters"** by recommending biomolecular methods to assess the impact of pollutants through the early detection of biodiversity and molecular shifts.

It also supports the following Key Enablers:

- **"Digital Ocean and water knowledge systems"** by recommending steps to improve digital infrastructure for biomolecular data, and their integration into Digital Twins of the Ocean.
- **"Public mobilisation and engagement"** by recommending Ocean literacy activities linked to biomolecular tools to help ensure that science-based marine policies are transparent and accessible to wider society.

The European Marine Board acknowledges that while the Working Group members (see Annex 1) who wrote this document and its recommendations represent diversity in terms of European geographical location, professional background, gender and career level, their views may not represent ideas from all forms of diversity. This document has a European focus, but its messages and recommendations are relevant to stakeholders globally.

Contribution to the UN Decade Challenges

This Future Science Brief and its recommendations support the UN Decade of Ocean Science for Sustainable Development's (Ocean Decade) challenges (C1 – C10) in the following ways:

- **"Understand and beat marine pollution" (C1)** by recommending biomolecular methods that can provide early-warning systems for "invisible" changes in marine life through the detection of pathogens, and molecular shifts associated with the impact of pollutants.
- **"Protect and restore ecosystems and biodiversity" (C2)** by recommending biomolecular methods for enhancing species inventories through the detection of rare, cryptic and invasive species, and for monitoring changes in biodiversity in response to human activities, climate change, and restoration.
- **"Sustainably nourish the global population" (C3)** by recommending biomolecular methods that can potentially inform fish stock assessments and be used for the early detection of harmful algal blooms and pathogens that can contaminate seafood and aquaculture.
- **"Develop a sustainable, resilient and equitable Ocean economy" (C4)** by recommending biomolecular methods for environmental impact assessments of industrial activities.
- **"Unlock Ocean-based solutions to climate change" (C5)** by recommending biomolecular methods for environmental impact assessments of marine carbon dioxide removal interventions and by increasing the incorporation of lower trophic[1] level dynamics into Earth system models.
- **"Increase community resilience to Ocean and coastal risks" (C6)** by recommending Ocean literacy and capacity development activities to foster increased engagement of citizens in biomolecular observations.
- **"Sustainably expand the Global Ocean Observing System" (C7)** by recommending innovation in low-cost, deployable *in situ* biomolecular samplers and sensors to collect biomolecular data at scale.
- **"Create a digital representation of the Ocean" (C8)** by recommending improved digital infrastructure for biomolecular data, and their integration into Digital Twins of the Ocean.
- **"Skills, knowledge, technology and participation for all" (C9)** by recommending skill development in biomolecular methods across a wide range of stakeholders.
- **"Restore society's relationship with the Ocean" (C10)** by recommending Ocean literacy activities linked to biomolecular tools to help ensure that science-based marine policies are transparent and accessible to wider society.

[1] Trophic refers to the feeding and nutrition relationships among organisms in an ecosystem.

1 Introduction

Marine biodiversity plays a key role in supporting human well-being through the provision of ecosystem services including food security, climate regulation, and recreation. Biodiversity also determines the resilience of marine ecosystems to disturbances and their capacity to adapt to environmental change, supporting long-term Ocean health (IPBES, 2019). This places the protection and sustainable use of biodiversity at the centre of global and European Ocean governance agendas e.g. the Marine Strategy Framework Directive (MSFD; Directive 2008/56/EC), the EU Biodiversity Strategy for 2030 (COM/2020/380 final), the European Ocean Pact (COM/2025/281 final), and the Agreement on Marine Biological Diversity of Areas Beyond National Jurisdiction (BBNJ Agreement; United Nations, 2023).

However, marine biodiversity is undergoing rapid and widespread change. From coastal ecosystems to the deep sea, the number of species is decreasing, community composition[2] is changing, and there is a loss of ecosystem integrity driven by multiple pressures that act simultaneously (United Nations, 2026). Human activities and climate change are reshaping the chemical and physical conditions of the Ocean with major consequences for marine biodiversity. Climate change is now the primary driver of change in marine ecosystems, affecting species distribution, physiological tolerance limits, and ecosystem function[3] through Ocean warming, acidification, marine heatwaves, and deoxygenation (Thangaraj & Sun, 2026). Fishing pressure, the exploitation of other marine resources (e.g. offshore energy), habitat modifications, and pollution also impact marine biodiversity, particularly in coastal zones where human activities are concentrated. Deep-sea ecosystems are poorly characterised and scarcely monitored, yet emerging and expanding uses of these ecosystems, such as for deep-sea mining and marine carbon dioxide removal (mCDR) are advancing rapidly and must be approached with caution (Sander et al., 2025).

There is a critical need for comprehensive assessments of marine biodiversity and ecosystems to understand how quickly and through which processes community structure and ecosystem functions are being altered in response to escalating and cumulative threats (Borja et al., 2024a). Our ability to understand what is changing and to respond effectively will depend on our capacity to observe marine life in all its dimensions: from microorganisms to megafauna, from surface waters to the deep sea, and from the presence of species to the functioning of the ecosystem. This requires reliable, rapid, and efficient tools that can detect biodiversity changes in both space and time. However, with up to two million marine species undiscovered (Bouchet et al., 2023), our ability to observe, understand, predict and ultimately respond to the impacts of human-induced pressures on biodiversity is challenged.

1.1. Traditional observation methods

Large-scale monitoring of marine ecosystems is essential for detecting early biological shifts. For centuries, marine scientists have used direct observation methods such as net sampling and visual, morphology-based species identification. These approaches have generated extensive knowledge of many marine ecosystems, including invaluable insights into life history and functional traits of organisms, and continue to be the base of most monitoring and observation programs. Traditional methods can also be combined with biomolecular tools for species descriptions through “integrative taxonomy” that provides a robust and holistic approach to the classification of marine diversity (Dayrat, 2005).

Morphology-based sampling and taxonomic expertise remain essential as they underpin the reference databases[4] on which the identification of molecules depends. They are also required to validate and interpret molecular results, and provide information on life stage, size, biomass, condition and behaviour that molecular signals alone cannot deliver.

However, complete species inventories using net capture and morphological identification can take years. Processing samples, counting and identifying organisms, and homogenising results from different observers is laborious and expensive and can include errors and biases (Sushmitha

[2] Community composition considers which species are present and in what abundance.

[3] Ecosystem function refers to the natural processes and interactions that take place within an ecosystem allowing it to sustain life e.g. nutrient cycling. Ecosystem services are dependent on a properly functioning ecosystem.

[4] Reference databases are curated digital libraries of information about biomolecules including their sequences, functions and the names of the organisms they belong to. They are used to compare unknown sequences against known sequences to identify their function and taxonomic origin.



Credit: Ruđer Bošković Institute, Center for Marine Research

Morphology-based sampling remains indispensable to underpin the reference databases on which molecular identification depends and to validate and interpret biomolecular results.



Credit: Ruđer Bošković Institute, Center for Marine Research

Some organisms are difficult to observe with traditional methods, and lose their shape during capture, such as this gelatinous salp.

et al., 2025). In addition, traditional methods provide a static snapshot (Thangaraj & Sun, 2026) of what species were present at a specific place and time, and often do not provide a comprehensive overview of biodiversity in large, complex ecosystems.

Direct observations face extreme challenges in the deep sea, which encompasses most of Earth's biome and remains underexplored. Various factors limit what can be physically sampled, including the remoteness of the open Ocean, the high costs of operating research vessels, extreme temperatures and pressures of the deep sea, and the spatial and temporal scales over which processes occur (Sushmitha et al., 2025). Our knowledge of deep-sea biodiversity is therefore fragmented and biased, with only a limited number of regions studied (see EMB Future Science Brief N°. 12 on on 'Deep Sea Research and Management Needs'; Sander et al., 2025).

Traditional methods are also biased toward certain types and life stages of larger organisms, since species identification based on morphology requires visible or easily identifiable characteristics. It is therefore most effective for fish and larger zooplankton, and is more challenging for microbes (Santi et al., 2021), small invertebrates and animals, and the larval, juvenile and resting stages of larger species. Microorganisms constitute at least two-thirds of marine biomass and are largely undetectable using traditional observation methods (Sushmitha et al., 2025). There are also animals that are difficult to observe, because they are nocturnal, cryptic[5], gelatinous, or rare (Yang et al., 2024).

[5] Cryptic species are morphologically indistinguishable from one another but are genetically distinct.

1.2. The biomolecular approach



Credit: © Nick Decombel

The emergence of biomolecular tools changes the way marine biodiversity is observed, contributing to a multidimensional understanding of ecosystem state and function.

Biomolecular methods were originally developed for medical diagnostics and human genomic research, and in the last two decades have been adopted in all fields of biology. Biomolecular tools can be used to study the genetic material (i.e. DNA and RNA), proteins and metabolites of marine organisms. Biomolecular methods include Polymerase Chain Reaction[6] (PCR)- based methods, and genomic[7], transcriptomic[8], proteomic[9], metabolomic[10] (collectively known as -omic[11]) technologies (see Chapter 3 for details). These can be used to analyse molecules from a single organism, from a multispecies sample (e.g. from a plankton tow), or from water and sediment samples (environmental DNA / eDNA[12]).

Biomolecular tools mitigate some of the biases of traditional observation methods and can fundamentally change the way we observe, monitor and assess marine biodiversity, adding value in terms of speed, non-destructive sampling, high sensitivity and resolution. They can detect a greater number of species and can be applied over broader spatial (open Ocean, deep sea) and temporal (seasonal and interannual) scales. Comprehensive molecular inventories capture the full spectrum of marine life, from microorganisms to megafauna, and from single cells to entire communities. -Omics approaches can reveal not only which species are present, but also what they are doing i.e. which metabolic pathways are active under particular conditions, how gene expression[13] changes in response to Ocean warming or acidification, and whether organisms possess and express genes that give them new capabilities (see Chapter 2). In this way, it is possible to better understand the resilience of ecosystems and begin to predict responses to future changes (Thangaraj & Sun, 2026). Biomolecular methods provide quantitative or semi-quantitative information on biomass, abundance, activity and function of organisms, contributing to a multidimensional understanding of ecosystem state and function (see Chapter 2).

High-throughput molecular sequencing technologies enable the simultaneous sequencing of many DNA molecules, resulting in the rapid generation of large datasets (see Guidi et al., 2020) through the characterisation of thousands of organisms from single water samples in weeks or days. Biomolecular tools are also increasingly cost-effective and accessible. High-throughput sequencing can cost as little as 100 euro per sample, portable sequencing systems can be used in the field, and standardised bioinformatic[14] tools and Artificial Intelligence (AI) systems increase the efficiency of sequence analysis. These developments lower the technical barriers, but specialist expertise continues to be essential for the design and validation of biomolecular analyses, bioinformatic choices, taxonomic identification, and the interpretation of uncertainty. A further advantage of biomolecular approaches is that they can be standardised and reproduced. Sampling, laboratory and bioinformatic workflows can be harmonised, updated over time, and documented in detail, and outputs are digital and therefore directly comparable, which allows integrated meta-analyses (see Chapter 5).

[6] PCR is a laboratory technique used to amplify copies of specific DNA sequences, generally present in small amounts, thus facilitating their detection and sequence analysis.
[7] Genomics is the study of DNA sequences, including genes, within a cell, tissue or organism (i.e. its genome) and their interactions with each other and the environment.
[8] Transcriptomics is the study of transcripts of a cell, tissue or organism across a variety of different biological conditions, which provides information on gene expression.
[9] Proteomics is the study of proteins produced by a single cell, tissue, organism, or ecosystem.
[10] Metabolomics is the study of the metabolites, or small molecules, present in an ecosystem, organism, cell or tissue under certain conditions.
[11] Omics refers to a suite of tools used to characterise and analyse entire sets of molecules (i.e, genes, transcripts, proteins, or metabolites) within an organism or environmental sample.
[12] Environmental DNA (eDNA) refers to genetic material found in the environment (e.g. water, sediment, soil, air) shed from organisms (e.g. cells, faeces) that can be used to study biodiversity without needing to capture or directly see the organism itself.
[13] Gene expression is a cellular process by which the instructions coded by DNA are converted into functional products, such as proteins. It can be investigated through transcriptomics.
[14] Bioinformatics is the use of computer science, mathematics and statistics to analyse large and complex biomolecular data.

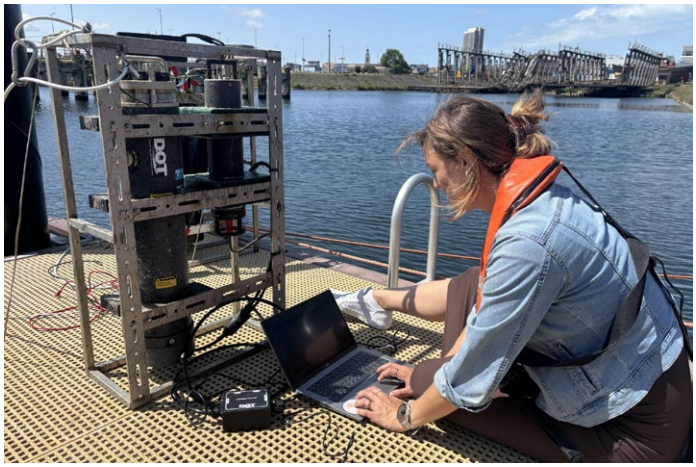
1.3. Application of the biomolecular approach to policy

Biomolecular methods for monitoring, managing, and protecting marine environments are currently used predominantly in research applications and need to be scaled up for use in routine monitoring. Global governance has aligned to create demand and opportunities to improve biological Ocean observations. The Kunming-Montreal Global Biodiversity Framework (CBD, 2022) has ambitious targets for monitoring and protection of biodiversity, which means genetic data need to be translated into information that managers can use while considering issues such as equity and global justice (Lehmann et al., 2026). The BBNJ Agreement has created a new governance structure for the high seas and explicitly recognises the need for environmental assessment tools, including biomolecular methods, to allow for conservation decisions on Marine Protected Areas (MPAs), Environmental Impact Assessments (EIAs), and to establish best practices for Access to Marine Genetic Resources (Humphries, 2025). These ambitions are enabled through the United Nations Decade of Ocean Science for Sustainable Development[15] (Ocean Decade), which demands improved observations to gain more knowledge on biodiversity and its role in marine ecosystems, along with better understanding and monitoring of marine ecosystems so they can be managed sustainably. The Global Ocean Observing System[16] (GOOS) is therefore advocating for the use of biomolecular technologies, and more collaboration between regions to expand observational capacity (Moltmann et al., 2019).

At the European level, several directives and strategies set legal requirements for extensive and comprehensive biodiversity and ecosystem monitoring, including the MSFD, the EU Biodiversity Strategy for 2030, the Habitats Directive (Council Directive 92/43/EEC); the Nature Restoration Regulation (Regulation (EU) 2024/1991); the Water Framework Directive (WFD, Directive 2000/60/EC); the Bathing Water Directive (BWD, Directive 2006/7/EC); the Common Fisheries Policy (CFP, (Regulation EU 1380/2013); the Ocean Pact and possible future Ocean Act (see Section 2.6.).

Simultaneously, there are European initiatives that are developing coordinated frameworks for biomolecular observations. The European Marine Omics Biodiversity Observation Network[17] (EMO BON; part of the Ocean Biomolecular Observing Network of the UN Ocean Decade structure) established and published protocols, standards and infrastructure that are now in operational use (see Chapter 5) to integrate molecular biodiversity data into the wider system of Ocean observation towards a European Ocean Observing System[18] (EOOS). In addition, as part of the Ocean Pact, the European Commission's OceanEye initiative (COM/2026/268 final) aims to improve governance of Ocean observations, and to support innovative markets for Ocean observation technologies, which will be critical for supporting innovation in biomolecular tools (see Chapter 4), and their wider uptake.

Marine observation systems have traditionally been dominated by physical and chemical measurements and need to better integrate both traditional and biomolecular biological observations (see EMB Future Science Brief N°. 3 on 'Strengthening Europe's capability in biological Ocean observations'; Benedetti-Cecchi et al., 2018). The integration of rapidly advancing biomolecular tools into an integrated (i.e. augmented) Ocean observation system will improve cost-effectiveness and increase accessibility to researchers. This must be done in a way that is scientifically reliable, operationally viable, and able to deliver the evidence required by existing legal and policy commitments, building on Horizon Europe projects such as BiOCEAN5D[19], Obama-Next[20] and GES4SEAS[21].



Credit: © Hanneloor Heynderickx

Support for innovation in biomolecular tools (e.g. eDNA samplers) is crucial for their wider uptake into an integrated Ocean observation system.

[15] https://oceandecade.org/
[16] https://goosocean.org/
[17] https://www.embrc.eu/emo-bon/
[18] https://www.eoos-ocean.eu/
[19] https://biocean5d.org/
[20] https://obama-next.eu/
[21] https://www.ges4seas.eu/

2 Applications for research, management and conservation

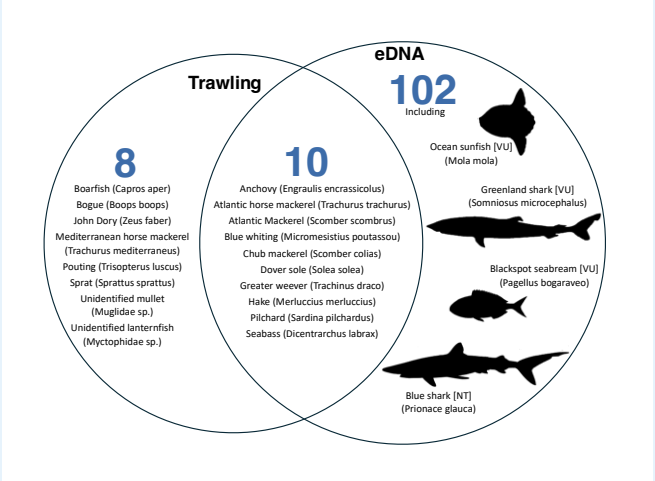
2.1. Summary

This chapter presents an overview of the implementation of biomolecular approaches in marine observation programmes and examples of potential policy applications (Figure 2.1.). It showcases the potential of these tools to provide insight into biodiversity and ecosystem structure and function across scales from molecules to the global Ocean (Zaiko et al., 2021).

2.2. "Who's there?" From presence to abundance in marine biodiversity assessments

The fundamental question of "Who's there?" underpins comprehensive marine biodiversity assessments. The ability to address this question has been advanced by the application of biomolecular techniques to monitoring large groups of marine microorganisms, revealing their previously unappreciated complexity (DeLong, 2009). In addition, some contemporary biomolecular methods have become generalised tools. For example, environmental DNA (eDNA) enables the non-destructive tracking of species presence, or recent presence, through analysing genetic material of larger marine organisms that is shed into the surrounding environment, fundamentally transforming the capacity to detect elusive, rare, cryptic, and non-indigenous species without direct observation or capture (see Box 2.1.). The speed and accuracy of biodiversity monitoring has been demonstrated by the European Autonomous Reef Monitoring Structures[22] (ARMS) project, which is part of the Marine Biodiversity Observation Network[23] (MBON) (ARMS-MBON) for benthic communities (Daraghmeh et al., 2025; see Box 3.3). Molecular techniques also enable quantitative measurements as proxies for organism abundance or biomass. For example, quantitative PCR[24] (qPCR; see Section 3.3.1.) allows targeted monitoring of species of conservation, economic (see Box 2.2.), or ecological significance.

Box 2.1. Metabarcoding to detect fish community composition



Credit: Modified from Fraija-Fernández et al. (2020) (CC BY 4.0).

Comparison of fish species caught in trawls and detected through eDNA metabarcoding. IUCN Red List categories: VU, Vulnerable; NT, Near Threatened.

In a study by Fraija-Fernández et al. (2020) in the Bay of Biscay, fish community composition was assessed by eDNA metabarcoding[25] from water samples, alongside trawling: the traditional method for surveying marine fauna. Both trawl and eDNA samples identified the same most abundant species, while trawling identified an additional eight species not found through eDNA, and eDNA detected 102 additional species, including species listed as threatened on the International Union for the Conservation of Nature (IUCN) Red List[26] (see figure). Metabarcoding of eDNA also yielded information on species distribution that aligned with expectations for the area, revealing different species occurrence patterns across depths. This study demonstrates the effectiveness of eDNA metabarcoding for potential application in the MSFD and CFP, including detection of the presence of rare and threatened species.

[22] https://www.vliz.be/projects/assembleplus/node/2326.html
[23] https://marinebon.org
[24] qPCR is a laboratory technique to amplify and simultaneously quantify a specific DNA or RNA marker in real-time.
[25] Metabarcoding is a technique that involves the PCR amplification and high-throughput sequencing of genetic marker regions from DNA extracted from mixed-community or environmental samples, enabling taxonomic characterisation of multiple organisms.
[26] https://www.iucnredlist.org/

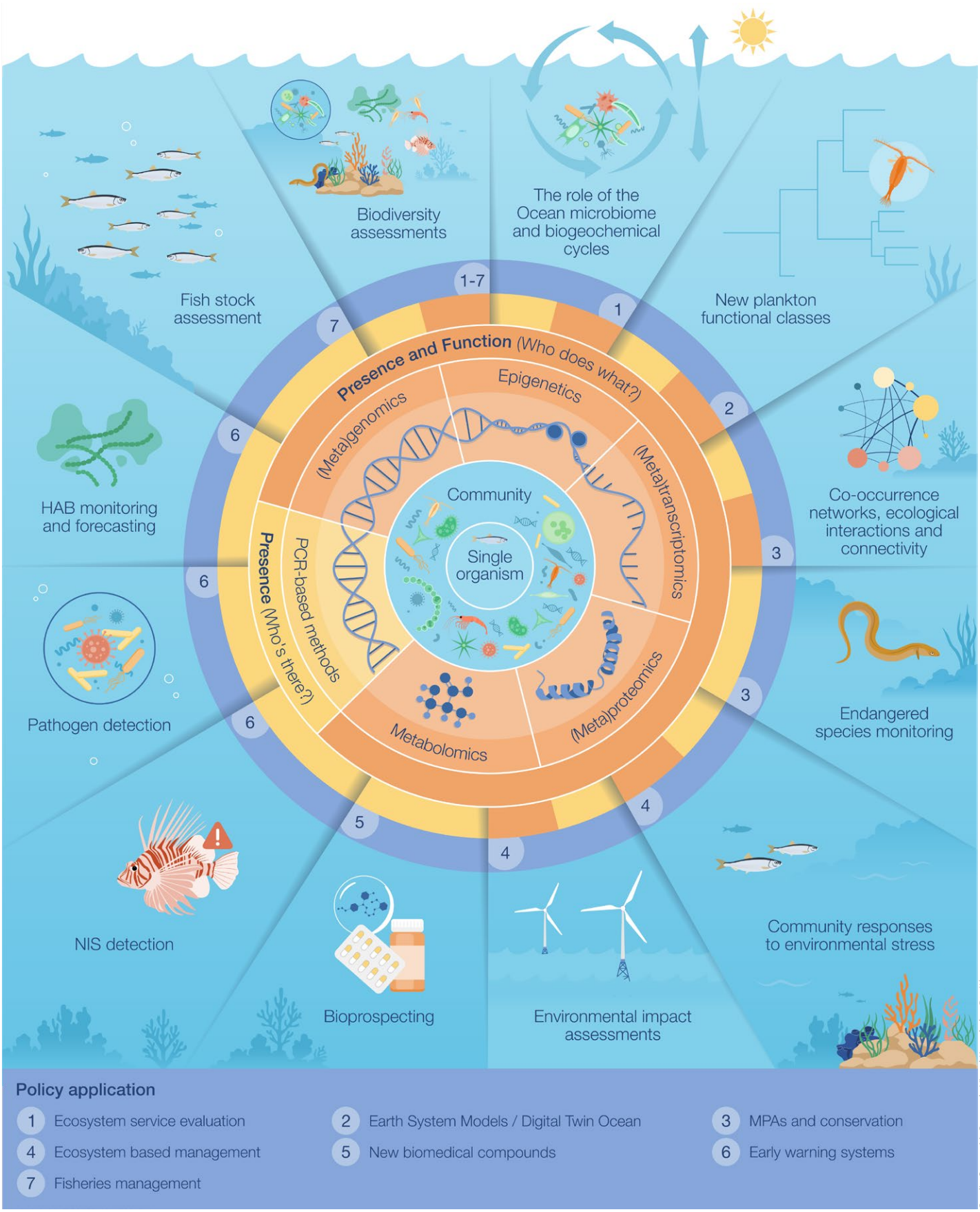
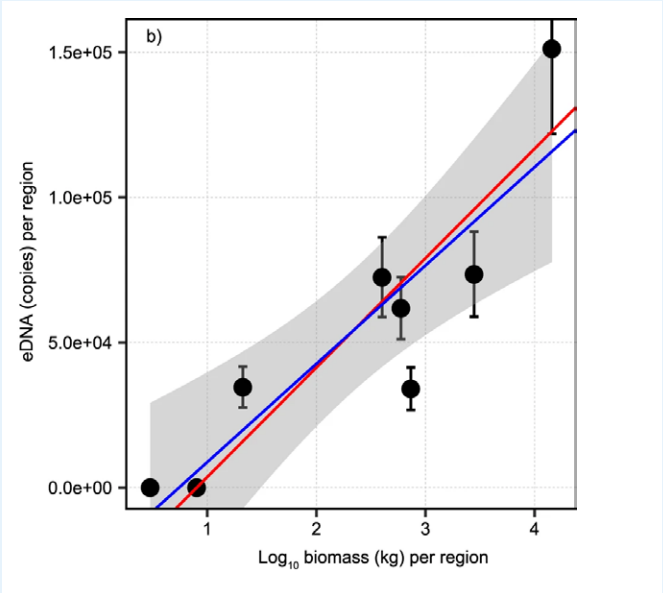


Figure 2.1. Overview of source material (either single organisms or communities), biomolecular methods, the scientific questions they can answer, and potential links to policy applications (indicated in numbers). Yellow indicates methods and applications linked to determining the presence of organisms (i.e. Who's there?), and orange indicates methods and applications linked to understanding presence combined with organism and ecosystem function (i.e. Who does what?). HAB = harmful algal bloom. NIS = non-indigenous species. MPA = marine protected area. PCR = polymerase chain reaction.

Box 2.2. Quantitative Polymerase Chain Reaction (qPCR) for fish stock estimation



Credit: Salter et al. (2019) (CC BY 4.0).

Correlation between Atlantic cod biomass obtained from trawl surveys and number of eDNA copies aggregated regionally.

A molecular assessment of Atlantic cod stocks in the North Atlantic by Salter et al. (2019) compared qPCR analysis from water samples with trawl-based detection of cod around the Faroe Islands. When data for each station were considered separately, qPCR generally detected cod whenever they were captured in the trawl, but there was not strong correlation between the quantities detected by the two techniques. However, at broader regional scales, eDNA concentrations directly correlated with fish biomass and Catch Per Unit Effort, an abundance measure used in fish stock assessment (see figure). Traditional fisheries management relies on trawl catches from multiple stations aggregated over larger spatial areas for fish stock assessments. Therefore, eDNA samples aggregated at a regional scale could be sufficiently sensitive to monitor stock fluctuations, minimising gaps in commercial fisheries data, and reducing the mortality of fish during the sampling process. Thus, eDNA monitoring potentially offers a powerful tool to support the CFP. A comparable study on Pacific hake similarly concluded that eDNA data aggregated at an appropriate spatial scale provided data suitable for management (Shelton et al., 2022). Broader ecological data, including on the diversity of fish and other organisms could be generated through applying metabarcoding (see Box 2.1. for example) to the same samples, in support of ecosystem approaches to fisheries management.

Biomolecular tools such as qPCR also enable surveillance of indicator species whose presence reflects ecosystem health or environmental conditions, including phytoplankton and bacterial populations, invasive species, harmful algal blooms (HABs), and pathogens that pose risks to human health and aquaculture operations. By facilitating detection at very low amounts and continuous monitoring, these methods strengthen early-warning systems and support timely, evidence-based response frameworks. qPCR can also be applied to eDNA-based detection of endangered species to provide actionable data for conservation planning with minimal disturbance to target organisms or their habitats. Digital PCR[27] (dPCR; see Section 3.3.1.) can provide highly sensitive detection and quantification of marine pathogens, microbes and viruses. For example, some toxic phytoplankton are present in low numbers before a bloom and dPCR allows their quantification before their numbers spike, helping to forecast blooms earlier and with greater accuracy (Lee et al., 2020), allowing timely adaptation of fishing or aquaculture activities.

Biomolecular approaches facilitate the discovery of new species across sizes from viruses to macro-organisms such as fungi or jellyfish. Global surveys such as those conducted by the Tara Oceans consortium[28] demonstrate an enormous reservoir of undescribed marine biodiversity, particularly among eukaryotes[29] and viruses (Sunagawa et al., 2020). The species discovery process has been enhanced through integrative taxonomy, which combines molecular and ecological data with traditional methods to define species boundaries with greater precision than morphology-based approaches alone, particularly for cryptic taxa[30] where molecular data are required for their identification. This is being enlarged through the monitoring undertaken by the BiOCEAN5D project.

Beyond species discovery, phylogenetic analyses[31] through multi-gene and whole-genome datasets have improved our understanding of evolutionary relationships that challenge traditional taxonomic frameworks. In addition, molecular taxonomic units such as operational taxonomic units[32],

[27] Digital PCR (dPCR) is a laboratory technique that amplifies and quantifies specific DNA segments.
[28] https://www.tara-oceans-science.org/
[29] A eukaryote is an organism whose cells contain a membrane-bound nucleus and other membrane-bound organelles, and includes animals, plants, and protists.
[30] Taxon (pl. Taxa) is a group of organisms that is classified together in taxonomy. Each taxon represents a unit at any rank in the taxonomic hierarchy, such as species, genus, family, order, class, phylum, or kingdom.
[31] Phylogenetic analyses are carried out to determine how closely related different organisms are.
[32] An operational taxonomic unit is a functional proxy for species when using DNA sequence data and is used to group organisms based on their genetic similarity, especially when it is not possible to identify every species precisely.

amplicon sequence variants[33] and functional gene markers complement classical Linnaean classifications[34], allowing researchers to assess fundamental ecological questions such as the drivers of biodiversity patterns and ecosystem resilience.

These same biomolecular tools can interrogate the past. The analysis of ancient DNA[35] (aDNA) extracted from marine sediment cores provides temporal snapshots of historical communities, enabling the reconstruction of past ecosystems and assessment of anthropogenic impacts, including pollution legacies, and the effects of climate change on marine biodiversity over decadal to millennial timescales (see Box 2.3.). Notably, the application of the highly sensitive dPCR technique (see Section 3.3.1.) to sedimentary aDNA has proven valuable for quantitatively tracing indicator species over time, enabling their use as proxies to reconstruct past environmental conditions. For example, sedimentary aDNA from microalgae associated with sea-ice has been used as a proxy to reconstruct past sea-ice conditions in Baffin Bay in Greenland (Harðardóttir et al., 2024).

Box 2.3. Sedimentary ancient DNA (aDNA) as a key to past biodiversity



Credit: Agnes Weiner.

Sedimentary ancient DNA can be collected using sediment cores.

Studying the impact of climate change and other pressures on marine biodiversity requires long-term observations to differentiate natural variability from anthropogenic-driven changes. Sedimentary aDNA allows retrospective analysis of changes in biodiversity, covering decades, centuries or millennia. Sedimentary aDNA is eDNA that has accumulated over time on the seafloor and preserved in the sediment, representing an archive for past biodiversity.

Once extracted, sedimentary aDNA can be analysed using metabarcoding or metagenomics[36]. The fragments extracted from the sediment are generally short because free DNA that exists outside of living cells degrades over time. Metagenomics in particular can retrieve information from very short, degraded fragments, and can verify whether DNA is indeed aDNA through assessment of the degraded ends (Armbrecht, 2020). Traditional paleontological approaches are limited to the few marine lineages with a fossil record (e.g. foraminifera[37] and diatoms[38]), but sedimentary aDNA analysis can determine the responses of entire past communities to past environmental change.

While still an emerging method, the analysis of sedimentary aDNA has already been successfully combined with traditional palaeoecological approaches to assess changes in biodiversity over the last 20,000 years in the Western Bering Sea in correlation with past changes in sea ice cover (Zimmermann et al., 2023). The study used metagenomics to reveal significant shifts in both benthic and pelagic communities during warmer, ice-free periods.

On shorter timescales, sedimentary aDNA has been used to study the impact of anthropogenic pollution on marine protists[39] over the last millennium in the Bay of Brest, France (Siano et al., 2021), revealing irreversible biodiversity shifts in response to pollution from World War II and contamination from 20th century agricultural practices. Sedimentary aDNA analysis holds great potential in the future, e.g. for tracing the arrival and potential settlement of invasive species, for establishing biodiversity baselines before anthropogenic impacts, and for long-term monitoring of the effectiveness of MPAs. In addition, the knowledge gained can be crucial for informing ecosystem and species distribution models to project future changes.

[33] Amplicon Sequence Variants are highly resolved identical sequences that allow precise identification of species.
[34] A Linnaean classifications use the traditional system of naming organisms into a ranked hierarchy based on shared physical characteristics e.g. kingdom, phylum, class, order, family, genus, species.
[35] Marine sedimentary ancient DNA (aDNA) is eDNA from organisms that lived in the Ocean in the past and that has accumulated over time on the seafloor and is preserved in sediment.
[36] Metagenomics is the study of genetic material obtained from organisms in an environmental sample (e.g. eDNA), enabling comprehensive identification of the organisms that are present, and characterisation of the collective genetic potential of the community.
[37] Foraminifera are single-celled marine eukaryotes, often with shells.
[38] Diatoms are microscopic, single-celled or colonial algae that have shells made of silica.
[39] Protists are a diverse group of eukaryotic single-celled (or colonial) organisms, such as unicellular algae, ciliates or amoebas, not classifiable as animals, plants, or fungi.

2.3. "Who does what?" From presence of organisms to marine molecular function assessments

Biomolecular approaches help to determine "Who does what?" in marine ecosystems. This allows biodiversity to be linked to ecosystem function by measuring functional processes, molecular mechanisms that underpin ecosystem services and impacts from environmental stress or exposure to toxins. Genomic and transcriptomic techniques are used to assess *in situ* genomic content and gene expression of natural populations under specific environmental conditions. The combination of these data with metabolic modelling (see Section 4.5.) indicates which organisms are potentially metabolically active and what biogeochemical processes they mediate.

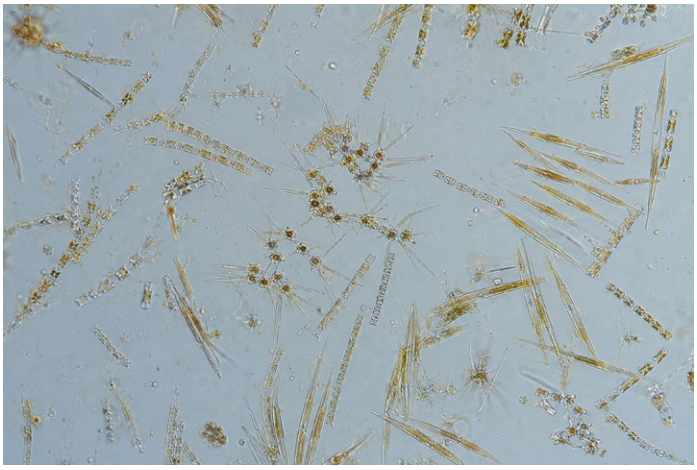
These techniques allow investigation across scales, from individual organisms to ecosystem-level applications, enabling the definition of marine biogeographical provinces (i.e. oceanic regions that show distinct biological signatures) based on suites of molecular functions and their expression patterns that reflect metabolic capacity, environmental adaptation, and biogeochemical potential across oceanographic gradients. These bioprovinces are based on comprehensive assessments of community composition, functional diversity, and ecosystem status, rather than on emblematic or charismatic species (Richter et al., 2022). An example of the use of this approach is in the assessment of the "Atlantification" of Arctic waters whereby water masses in the Arctic change to show properties associated with the Northern Atlantic Ocean.

Holistic biodiversity inventories capture the full taxonomic and functional spectrum of life from viruses to vertebrates, and this can support the development of new approaches to the designation and monitoring of MPAs (Esteban-Cantillo et al., 2024) and other area-based management tools that protect the microbial and meiofaunal[40] diversity that sustains marine ecosystem services (e.g. European Marine Board, 2019). For example, the larvae of coral and fish can be tracked using species-specific qPCR assays on eDNA from water masses. Their presence can be aligned with data on oceanic currents to identify source and sink populations and to validate oceanographic connectivity models. This informs the spacing and sizing of MPAs, identifying priority source reefs, and validating larval export between protected and unprotected areas (Underwood et al., 2018).

This functional perspective has also facilitated the identification of new plankton functional types by characterising their metabolic capabilities, physical properties, nutrient requirements, and biogeochemical roles. These new plankton functional types are being integrated into Earth System Models to improve predictions of Ocean productivity, carbon cycling, nutrient regeneration, and climate feedback (Régimbeau et al., 2025). Metabolically defined functional groups recognise that organisms that are distantly related may perform similar ecological functions while those that are closely related may occupy distinct functional niches. The integration of new plankton functional types into global biogeochemical models advances our understanding of how marine ecosystems will respond to environmental change.

Biomolecular approaches enable the prediction of species interactions by identifying metabolic cross-talk[41]. For example, genomic and metagenomic analyses have refined our understanding of vitamin dependencies in marine food webs including auxotrophy[42], trophic interactions, and community assembly in ways that are invisible using traditional methods. These vitamin dependencies create partnerships that shape community composition, influence competitive outcomes, and generate networks that determine ecosystem stability and resilience. This has fundamentally improved our understanding of microbial food webs and the mechanisms maintaining phytoplankton diversity (Giordano et al., 2024).

Functional insights from biomolecular data also refine understanding of microbial engines (Falkowski et al., 2008) and the recognition that specific groups of microbes with similar ecological functions drive key transformations in the nitrogen, sulphur, phosphorus, and carbon cycles. Metagenomic and metatranscriptomic[43] provide baselines for understanding the regulatory mechanisms involved in nutrient cycling and quantitative prediction of responses to pressures including warming, acidification, and deoxygenation. This enables genomically informed models of Ocean processes that include microbial diversity, functional, and environmental regulation of gene expression.



Credit: University Urbino, CoNISMA.

Biomolecular approaches allow the identification of new plankton functional types that are being integrated into Earth System Models.

[40] Meiofauna are small invertebrates living in aquatic sediments.

[41] Metabolic cross-talk is biochemical communication between organisms and cells which influences function, survival and growth.

[42] Auxotrophy is the inability of an organism to synthesise a particular compound required for its growth and survival, which it must therefore obtain from its environment.

[43] Metatranscriptomics is the study of which genes are being actively expressed (e.g. from mRNA from organisms in an environmental sample), showing which metabolic pathways are active under specific environmental conditions.

Box 2.4. Novel sources and diversity of natural products from the global Ocean microbiome



Credit: Myriam Zilles, Unsplash

Genomes and metagenomes can be screened for biosynthetic genes, enabling a more efficient discovery pipeline for pharmaceuticals, anticancer agents and antimicrobial compounds.

The discovery of natural products (i.e. compounds from nature that are useful in food or medicine) traditionally involves screening species for functions of interest, then interrogating the genes coding for those functions. Over time, gene clusters have been described for a vast array of natural product biosynthetic pathways[44], and now genomes and metagenomes can be screened for similar genes. This approach enables the isolation of target organisms, reducing the time and costs associated with screening large numbers of specimens. It also facilitates a more efficient discovery pipeline for pharmaceuticals, anticancer agents and antimicrobial compounds. For example, the screening of 35,000 metagenomes and single-cell genomes from the global Ocean microbiome revealed 40,000 different biosynthetic gene clusters, most of which were novel (Paoli et al., 2022). Most groups of organisms with the highest number of biosynthetic gene clusters were previously not known to possess biosynthetic activity. The dominant taxonomic group producing natural products with high biosynthetic potential are bacteria that are as yet uncultivated. The data from this study were made available to the public for exploration via an interactive web interface: "Ocean Microbiomics Database"[45].

Functional genomics can also be used for bioprospecting[46] (see Box 2.4.). The global antimicrobial resistance crisis has accelerated the search for new antibiotics from marine microbial sources, which harbour vast reservoirs of genes encoding previously unknown molecules. Marine organisms, particularly bacteria, archaea[47] and fungi, produce diverse metabolites as chemical defences and signalling molecules, many of which exhibit potent antimicrobial or anticancer activity. Genomic mining[48] enables the discovery of these genes directly from eDNA, bypassing the cultivation bottleneck that has historically limited access to marine microbial chemistry.

Curiosity-driven research into the ecological roles of microbial processes can generate technologies with impacts far beyond their original context, reinforcing the imperative for sustained investment in fundamental marine science. For example, the CRISPR-Cas[49] system (see Section 3.3.1.3.), which is revolutionising molecular biology, biomedicine, and agricultural biotechnology, originated from studies of bacterial adaptive immunity in marine and terrestrial microbiomes. Similarly, some deep-sea microbes harbour thermostable enzymes that are crucial for PCR reactions, underscoring the profound value of ecosystem-level investigations for both fundamental understanding and practical applications.

Bioprospecting efforts are increasingly focused on identifying microbial processes relevant to marine carbon dioxide removal (mCDR) approaches, including enhanced biological carbon pumps, mineral carbonation pathways, and engineered phytoplankton with augmented carbon fixation or export efficiency (Belkin et al., 2026). Geoengineering applications demand critical caution (see EMB Future Science Brief N°. 13 on "Monitoring, Reporting and Verification for Marine Carbon Dioxide Removal"; Muri et al., 2025), and molecular approaches are essential to understand their ecological consequences, such as trophic cascades, biogeochemical perturbations, shifts in community composition, and unintended feedback on climate systems.

Functional genomics enables a new generation of ecosystem modelling by integrating metabolic profiles of organisms. When combined with environmental data, this allows genome-scale metabolic reconstructions to predict organism-level and community-level responses to environmental change (Régimbeau et al., 2025). These models simulate cellular metabolism under varying nutrient availability, temperature regimes, and light conditions, generating testable predictions about growth rates, resource competition, and ecological interactions that can be validated through laboratory experiments and field observations.

[44] Biosynthetic pathways are a series of chemical processes that build complex molecules.
[45] https://microbiomics.io/ocean/
[46] Bioprospecting refers to searching for, discovering, and utilising biological resources such as genes, enzymes, or other molecules from nature for commercial or scientific purposes.
[47] Archaea are single-celled prokaryotic microorganisms that are widespread in (extreme) marine and terrestrial environments.
[48] Genomic mining is the process of searching through genomes to identify genes that may have interesting or useful functions.
[49] CRISPR-Cas is a tool to edit DNA with high precision, using a system from the natural defences of bacteria.

Finally, biomolecular approaches provide unparalleled insights into evolutionary processes by identifying and quantifying genetic mutations, which provide real-time observations of adaptive evolution in response to environmental stressors such as Ocean acidification, warming, pollution, and hypoxia. By sequencing populations across environmental gradients or time, researchers can detect beneficial genetic mutations in populations, versions of genes that provide an advantage, and quantify rates of evolutionary change, thereby describing the evolutionary potential and constraints that will determine marine biodiversity trajectories under global change scenarios. This capacity to observe evolution in action can transform our collective ability to predict and manage biological responses to anthropogenic environmental change.

2.4. Investigating the Ocean across scales

Molecular techniques can be used to holistically investigate marine systems across a range of spatial and biological scales, from individual molecules to the global Ocean, covering approximately twelve orders of magnitude (Figure 2.2.). These tools provide common currencies, from DNA sequences, gene expression profiles, metabolite concentrations, and protein abundances, that can be measured consistently with reproducible metrics across scales. This enables mechanistic linkages between molecular processes occurring within individual cells and emergent properties that manifest at ecosystem and global scales, such as primary productivity, nutrient cycling, carbon export, and climate regulation.

At the finest scales, genomics and metabolomics describe the molecular machinery that determines organism physiology, metabolic capabilities, and environmental responses, whilst proteomics and transcriptomics capture the dynamic regulation of these systems under varying environmental conditions. Scaling upward, molecular data illuminate the organisation and biological functions of individual organisms, including their acclimation via physiological plasticity[50].

Fine-scale biomolecular applications include their use in defining the concept of the holobiont, as an ecological unit comprising a host organism and its associated microbial

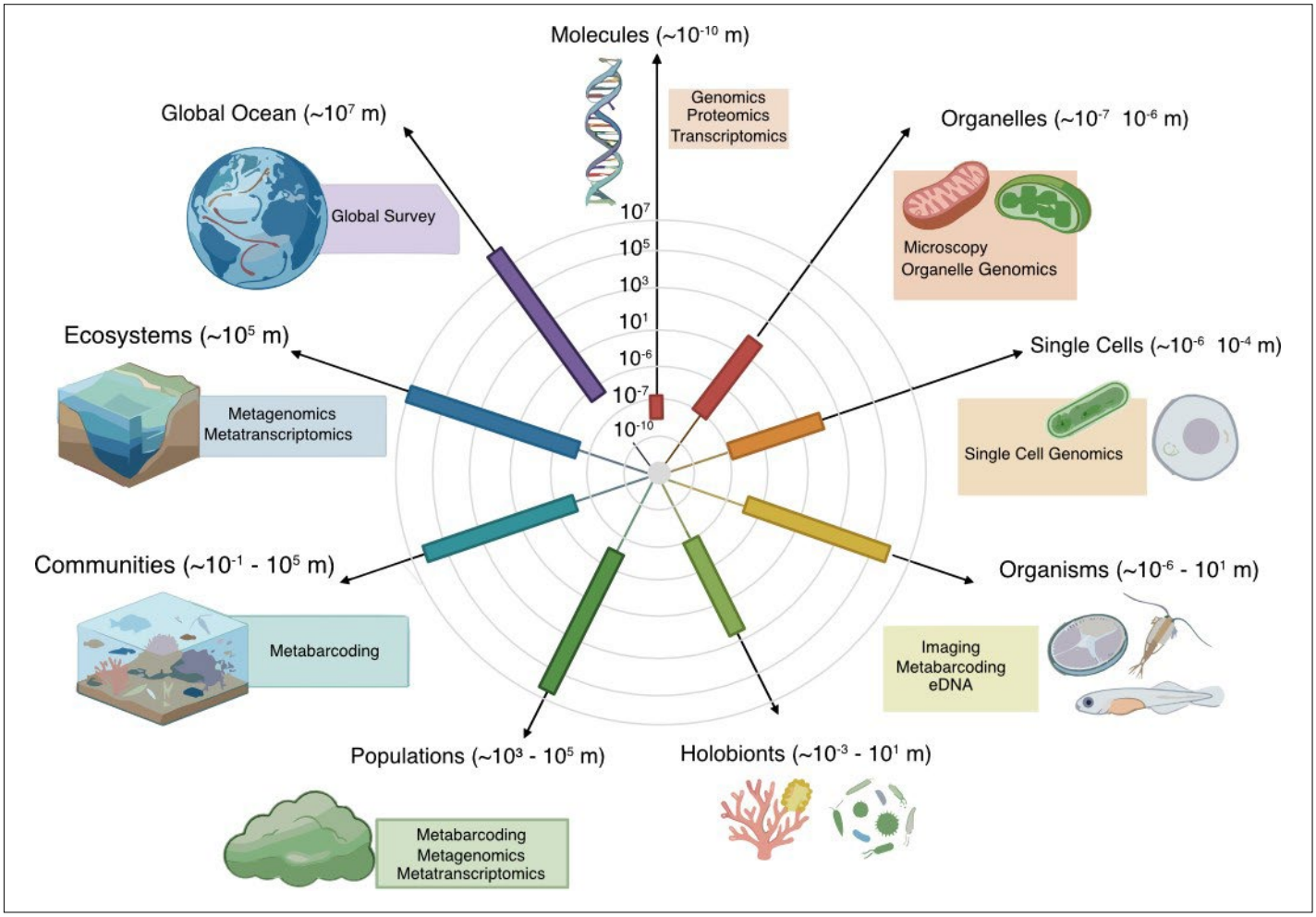


Figure 2.2. Illustration of spatial and organisational scales of molecular analyses, from individual molecules to the global Ocean, encompassing approximately twelve orders of magnitude. Coloured bars show the scale with which the case study is associated in metres.

[50] Physiological plasticity is the ability to adapt functions and processes in response to changes in the environment.

communities (Dittami et al., 2021). This recognises that multicellular organisms exist as complex assemblages of interacting genomes, with microbial symbionts contributing essential functions including nutrient acquisition, immune defence, developmental regulation, and environmental stress tolerance. In marine systems, this has led to the definition of the microenvironment surrounding plankton cells, known as the phycosphere, and the broader concept of the microbiome to describe the microbial consortia associated with animal, plant, and algal hosts.

Molecular approaches can reveal cryptic biological phenomena that are not easy to detect and that fundamentally reshape our understanding of cellular evolution and marine ecosystem functioning. For example, the understanding of symbiotic integration in marine ecosystems has been advanced using both proteomics and microscopic analyses to identify a nitroplast (i.e. a nitrogen-fixing structure) in marine algae, which originated from an endosymbiotic cyanobacterium[51] that was engulfed by algae (Coale et al., 2024). The nitroplast enables the host cell to incorporate biologically fixed nitrogen under conditions where dissolved inorganic nitrogen is limited, contributing to primary productivity and supporting food webs in nutrient-poor waters. At community and ecosystem levels, -omics data enable the identification of co-occurring species and the inference of ecological interactions through network analyses. Co-occurrence and association networks identify statistically structured relationships among taxa from metagenomic or metabarcoding data. These associations may reflect direct interactions, shared environmental responses, dispersal or other indirect processes, and therefore require experimental, temporal, spatial or functional validation before being interpreted as ecological relationships such as mutualism[52], commensalism[53], competition, or shared environmental preferences (Figure 2.3.). When integrated with functional genomic data, these co-occurrence networks can be interpreted mechanistically, creating hypothesis-driven models of species interactions (Lambert et al., 2024). For example, the identification of bacteriophage[54] sequences and their hosts underpins predator-prey dynamics that regulate microbial population densities. Co-occurring species also help to understand gene transfer between unrelated organisms. These -omics-informed models enhance our capacity to assess microbial ecosystem structure, predict community responses to environmental perturbations, and identify taxa whose functional roles disproportionately influence ecosystem stability and resilience.

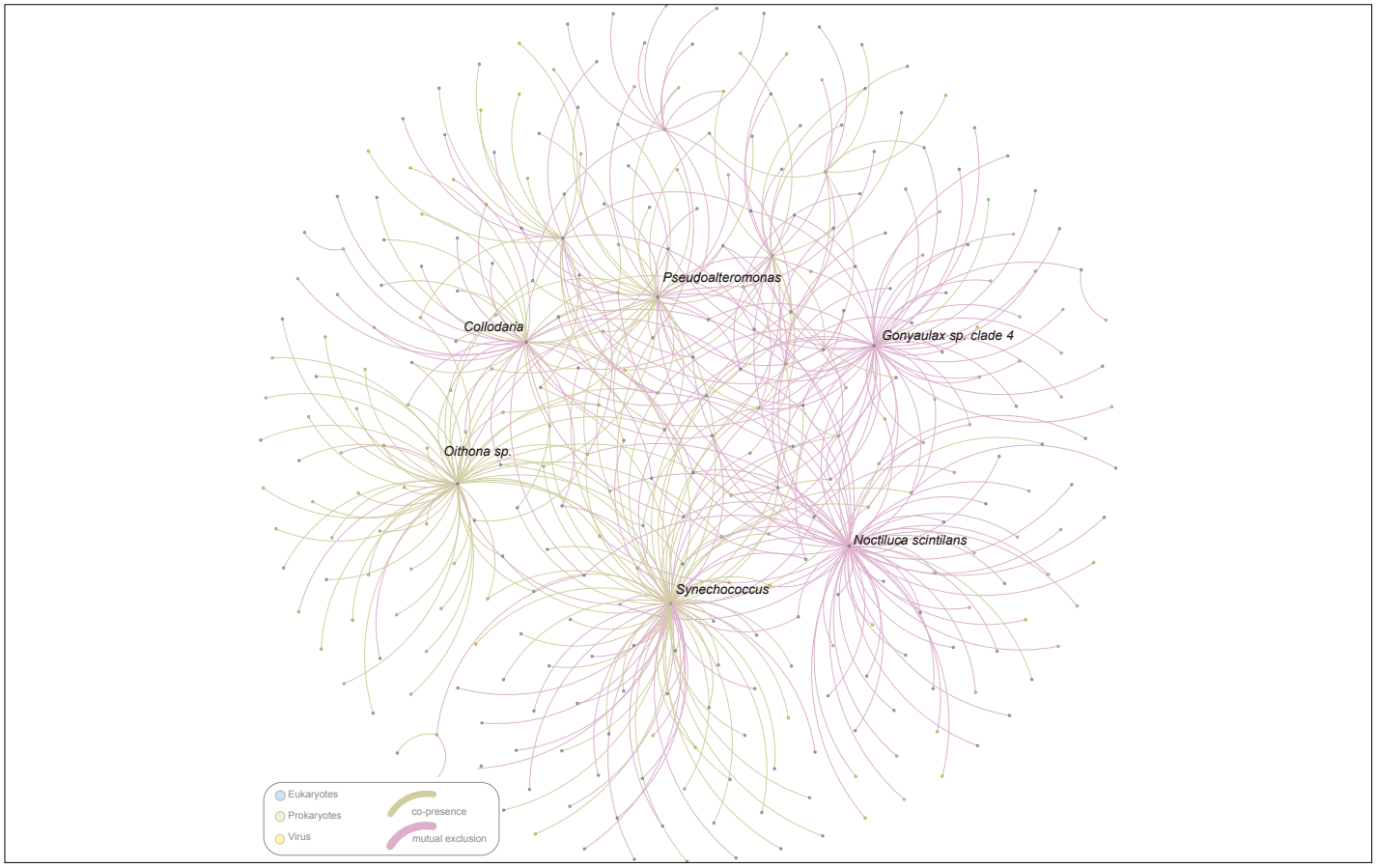


Figure 2.3. Co-occurrence of plankton organisms associated with carbon export in the oligotrophic (i.e. nutrient poor) Ocean, based on network analysis of the Tara Oceans dataset. Nodes are organisms (the most emblematic species are highlighted by their Latin names). The lines represent significant co-occurrence between organisms (positive depicted in green lines and negative in purple lines).

[51] Endosymbiotic cyanobacterium is a type of microscopic, photosynthetic bacteria that lives inside another organism in a symbiotic relationship.
[52] Mutualism is a type of relationship between two different species in which both benefit from the interaction.
[53] Commensalism is a type of relationship between two different species in which one benefits while the other is neither helped nor harmed.
[54] A bacteriophage is a type of virus that infects and replicates inside bacteria.

At the largest scales, following the Sorcerer II expedition (Yooseph et al., 2007), global Ocean surveys such as Tara Oceans (Box 2.5.) deployed standardised biomolecular protocols across oceanographic provinces, generating global datasets that show biogeographical patterns, environmental drivers of microbial community composition, and the functional genomic architecture of the global Ocean microbiome. These efforts link molecular mechanisms to planetary biogeochemical cycles and climate systems, providing holistic insights across the full spectrum of planktonic organisation. For example, through repeated sampling and sequencing of populations within water masses, we can observe the evolution of biological communities through space and time as the water mass is transported. This four-dimensional perspective (3D space and time) reveals how selection pressures, gene flow, and genetic drift[55] shape population structure outside of traditional biogeographic boundaries (Richter et al., 2022). For example, population genomic analyses of species sampled along Ocean currents shows rapid adaptive responses to temperature gradients, providing insights into evolution in pelagic ecosystems (Guérin et al., 2022). These insights challenge classical ecological theories developed primarily from terrestrial systems, demonstrating that marine populations are much more connected and adaptable than previously recognised (Richter et al., 2022). Understanding these evolutionary dynamics is essential for predicting species' capacity to adapt to climate change, as genetic variation and mutation rates determine their adaptive potential.

Box 2.5. Uncovering functional ecosystem processes in the Ocean with metagenomics and metatranscriptomics



Credit: © Louise Cognard, Fondation Tara Ocean

The global Tara Oceans Expedition transformed scientific understanding of marine ecosystem function through -omics techniques.

The global Tara Oceans Expedition transformed scientific understanding of marine ecosystem function using metagenomics and metatranscriptomics. Between 2009 and 2013 the expedition collected thousands of samples from plankton communities across the world's Ocean, spanning viruses, bacteria, protists, and small animals. By sequencing DNA and RNA directly from environmental samples, researchers identified which organisms were present, determined which genes were active and what ecological functions they performed (Sunagawa et al., 2015). They showed how microscopic organisms collectively regulate critical ecosystem processes such as carbon fixation, nutrient cycling, and food-web interactions. Metagenomic analyses uncovered an enormous diversity of previously unknown microbial genes involved in nitrogen and phosphorus metabolism, indicating that microbial communities play a far greater role in nutrient transformations than previously recognised. Complementary metatranscriptomic analyses showed that gene activity changes rapidly in response to environmental conditions, such as temperature shifts or nutrient availability, linking microbial metabolism directly to oceanographic dynamics.

Identifying genes and metabolic pathways that are active in different regions and conditions allows inference of how marine ecosystems respond to climate variability, pollution, or other pressures. Such knowledge supports ecosystem-based management by improving mechanistic understanding of primary productivity and nutrient cycling, which is increasingly integrated into biogeochemical and ecosystem models used to inform food-web dynamics, fisheries management, carbon budgeting, and biodiversity monitoring. These models can provide early-warning systems for adaptive management and conservation planning. The Tara Oceans initiative demonstrates how biomolecular approaches can provide a powerful framework for monitoring Ocean health. Integrating metagenomic and metatranscriptomic observations into long-term Ocean monitoring programs could enable early detection of ecosystem shifts and improve the scientific basis for conservation and sustainable management of marine resources.

[55] Genetic drift is the change in frequency of a gene variant in a population from generation to generation due to chance, rather than natural selection.

2.5. Monitoring the Ocean and predicting its future

High-throughput molecular technologies together with advanced modelling allows data to be produced at large spatial scales and at high resolution. The following subsections provide examples of how these biomolecular observations and modelling can be applied to Ocean monitoring and prediction.

2.5.1. Molecular technologies for the sustainable use of our Ocean

Highly-sensitive -omics techniques can be employed to enable the sustainable use of our Ocean across industries. eDNA tools are increasingly employed in monitoring programmes across industries. Gene expression provides real-time data to measure baseline biodiversity versus the impact of industrial activities on Ocean biodiversity, highlighting known and unknown Ocean stressors, such as toxic metals or oxygen depletion. They can be used to understand biological impacts across the life cycle of offshore energy exploration, production, and decommissioning (e.g. Environmental Genomics Research Programme, 2025). In addition, biomolecular methods have been used to understand the ecological impact of the Deepwater Horizon oil spill in the Gulf of Mexico (e.g. Mason et al., 2012) and to measure distinct metazoan biodiversity across polymetallic nodule[56] fields (Laroche et al., 2020). Metabarcoding can complement Environmental Impact Assessments (EIAs) by identifying species that are difficult to identify with imaging such as meiofauna and microbes and by providing in-depth understanding of ecological functions and interactions. Metabarcoding can also deliver species inventories over time to complement those collected via imaging surveys and collections (Rabone et al., 2023) to understand background ecological dynamics and effects of industrial activities on specific taxa. State-of-the-art methods and opportunities for implementing environmental genomics are discussed by marine industries and survey services in workshops such as the International Workshop on Environmental Genomics[57]. This will be important for the EIAs required for activities within Areas Beyond National Jurisdiction as part of the BBNJ Agreement.

2.5.2. From reactive surveillance to predictive pathogen tracking

Biomolecular tools have advanced pathogen surveillance in aquaculture and wild marine populations, shifting from reactive outbreak management to proactive risk assessment. qPCR and metagenomics enable early detection of viral, bacterial, and parasitic pathogens before clinical symptoms manifest, providing critical lead time for intervention (e.g. Benedicenti et al., 2026). These tools can trace pathogen dispersal across Ocean basins by combining eDNA detection with Ocean circulation models to map pathogen transmission routes, identify source populations, and predict downstream infection risks (Drouet et al., 2021). This is particularly crucial for managing transboundary disease threats, such as the spread of infectious salmon anaemia virus or oyster herpes virus, where understanding dispersal pathways informs coordinated international biosecurity responses and enables anticipatory management strategies to prevent spread into economically vital aquaculture systems.



Credit: Jesse de Meulenaere, Unsplash.

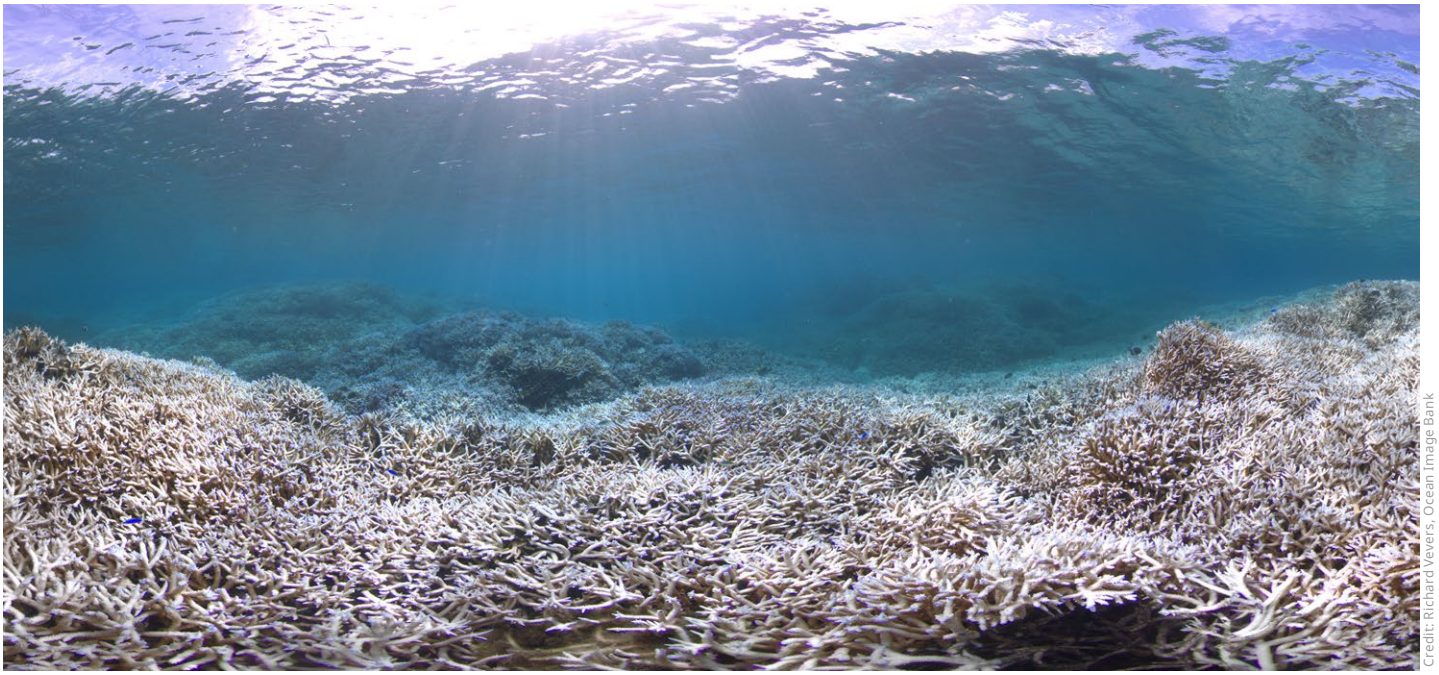
Biomolecular tools can be used to understand biological impacts across the life cycle of offshore energy.

2.5.3. Anticipating critical transitions and tipping points

The molecular characterisation of marine communities could provide early-warning signals of ecosystem tipping points, which are critical thresholds beyond which systems undergo rapid, often irreversible state changes. Time-series metabarcoding and metatranscriptomic data show shifts in community composition, functional gene expression, and metabolic activity that precede observable ecosystem collapse. For example, molecular monitoring of microbial communities in coral reef systems identified early indicators of microbial imbalances that predict bleaching events before signs of visible stress (Wegley et al., 2018), which opens opportunities for *in situ* molecular monitoring. In addition, changes in phytoplankton community structure detected using biomolecular methods can signal impending regime shifts in pelagic ecosystems (Wollschläger et al., 2014).

This anticipatory capacity has the potential to support adaptive management by enabling earlier detection of ecosystem change and informing intervention before irreversible damage occurs. However, the reliable prediction of ecosystem tipping points from molecular indicators remains an emerging field, and current evidence is not yet sufficient to establish generalisable molecular thresholds or reliable forecasts of when critical transitions will occur. Integrating molecular time-series data with environmental observations and dynamic ecosystem models may improve this predictive capacity, but substantial research is needed before they can be used in ecosystem management and policy.

[56] Polymetallic nodules are rocks on the deep-sea floor that contain natural deposits of minerals including manganese and iron.
[57] https://iweg.global/



Credit: Richard Vevers, Ocean Image Bank

Molecular monitoring of microbial communities can identify early indicators of microbial imbalances that predict coral bleaching events.

2.5.4. Assessing acclimation and adaptation under environmental change

Molecular tools enable mechanistic predictions of how marine organisms will respond to multiple climate stressors, including Ocean warming, acidification, expanding oxygen minimum zones, and microplastic pollution. Transcriptomic and proteomic analyses reveal the physiological mechanisms underlying stress tolerance and identify the molecular pathways activated during acclimation. Quantifying gene expression responses across life stages and environmental gradients allows prediction of critical thresholds beyond which acclimation capacity is exceeded. For fast-growing species, experimental evolution studies can distinguish acclimation from heritable genetic adaptations, informing projections of the potential to rapidly develop new traits (Brennan et al., 2022). Importantly, biomolecular approaches determine synergistic effects of multiple stressors. For example, exposure of zooplankton to microplastics under hypoxia conditions induces oxidative stress and leads to antioxidant depletion, which reduces their resilience to low oxygen conditions (Lee et al., 2022). These mechanistic insights can enable more realistic parameterisation of ecosystem models (i.e. Digital Twins of the Ocean (DTOs)) including -omics; see Section 4.5.) that account for organism-level responses propagating through food webs, ultimately determining ecosystem resilience and the maintenance of critical functions under environmental change.

2.5.5. Redefining ecosystem services through functional genomics

The molecular characterisation of the functional roles of marine organisms may require an adaptation to the current definition of ecosystem services (see EMB Future Science Brief N°. 5 “Valuing Marine Ecosystems: Taking into account the value of ecosystem benefits in the Blue Economy”; Austen et al., 2019). Functional genomics shows that ecosystem functions are encoded in the collective metabolic potential of communities (Guidi et al., 2016). This means that the analysis of metagenomics allows us to quantify genes and how they are associated to each other generate functions that are responsible for nutrient cycling, carbon sequestration, pollutant degradation, and secondary metabolite[58] production, which should be included in the valuation of ecosystem services. For example, marine microbes are typically valued for their role in primary production and as proxies for assessing environmental stressors. However, they also contain a variety of antibiotic biosynthetic pathways, meaning that they hold additional value with regard to their potential for pharmaceutical applications (Martiny et al., 2026). Similarly, quantifying the genomic capacity for nitrogen fixation or methane oxidation[59] provides mechanistic links between biodiversity and biogeochemical services. Finally, fusion of gene and satellite data can be used to assess Ocean provinces associated with carbon export to the seafloor, net primary production or gene diversity (Serandour et al., 2026). Thus, ecosystem services should be redefined by integrating -omics functional descriptions.

[58] Secondary metabolites are chemicals produced by organisms that are not directly involved in essential processes but have functions such as defence, competition or communication.

[59] Methane oxidation is a process which reduces methane in the environment, thereby reducing warming.

2.6. From molecular metrics to policy implementation

To achieve policy objectives, biomolecular methods for monitoring, managing, and protecting marine environments need to be scaled up from research applications to routine monitoring, and integration into adaptive management. Molecular methods are evolving from complementary research tools into operational components that can support risk assessment, ecosystem-based management, European policy implementation, and regional and international frameworks (Leese et al., 2016). This requires translating complex genomic data into interpretable indices that inform management decisions (Pawlowski et al., 2018) e.g. the development of molecular ecosystem health indices, pathogen risk scores, and functional diversity metrics that complement traditional indicators (Cordier et al., 2021). Bioinformatic products (i.e. outputs or tools created by the computational analysis of biomolecular data) need to be incorporated into DTOs, such as the European EDITO[60]. This requires investments in data repositories and visual interfaces that translate complex genomic data into an open science framework (see Chapter 5).

The European Union has funded the development of a scientific and technical framework to integrate molecular methods into routine environmental monitoring through initiatives such as DNAqua-Net[61], Genomic Observatories Metabarcoding for Environmental Assessment of European Seas[62] (GEANS), BIOcean5D, eDNAqua-Plan[63] OBAMA-NEXT, MBON and eDNA4Policy[64]. These initiatives are working to harmonise molecular protocols, establish reference databases (see Section 3.3.3.), develop quality assurance and quality control procedures, and evaluate DNA-based indicators for regulatory applications. They aim to improve the comparability of molecular data across laboratories and countries, while providing the scientific and methodological basis for evaluating the potential integration of molecular approaches into monitoring frameworks associated with European marine policies.

Rather than replacing conventional monitoring, these projects are developing molecular approaches as complementary tools capable of enhancing biodiversity assessment, early detection of ecological change, and evidence-based ecosystem management. Collectively, these initiatives represent an important transition from proof-of-concept studies towards operational implementation, providing the scientific basis for the future incorporation of molecular methods into European marine monitoring programmes.

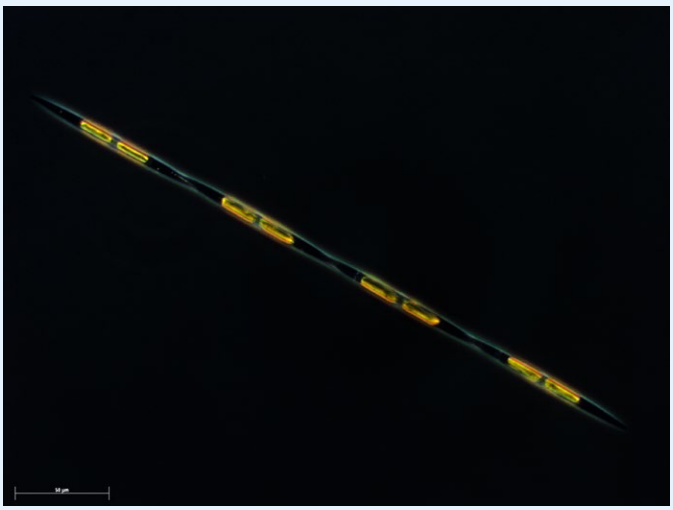
Although the Global Ocean Observing System (GOOS) Essential Ocean Variable[65] (EOV)descriptors embrace molecular tools, more work is needed to fully implement molecular techniques into Essential Biodiversity Variables[66] (EBVs). The Group on Earth Observations Biodiversity Observation Network[67] (GeoBON) established six classes of EBVs: Genetic Composition, Species Populations, Species Traits, Community Composition, Ecosystem Functioning, and Ecosystem Structure. Within Genetic Composition, metrics based on biomolecular data exist for populations, and eDNA data could generate the metrics required for Species Populations, but genomics remains underutilised across the other EBV classes (Lumbierres et al., 2024), even though molecular data could provide information on these classes. EBVs require large spatial and temporal coverage and eDNA can provide standardised and scalable approaches to monitoring under European Directives such as the MSFD. However, this will require the development of bioinformatically-derived policy-specific indicators for more efficient, accurate and cost-effective monitoring.

Recent evaluations of the MSFD (COM/2025/3 final) have highlighted that the assessment of Good Environmental Status (GES) remains inconsistent across descriptors, regional seas and Member States (Nikolaou et al., 2025). Consequently, there is growing recognition that harmonised and innovative monitoring approaches, including molecular methods, are needed to strengthen MSFD implementation (Palialexis et al., 2021). For example, combining metabarcoding of stomach contents with other dietary information to inform who eats whom, can contribute to D4: food webs. qPCR and dPCR are increasingly being trialled as complementary approaches to traditional methods for supporting assessments under MSFD descriptors, including D1: biodiversity, D2: non-indigenous species, D3: commercial fish and shellfish, D4: food webs, and D5: eutrophication. However, the routine implementation of biomolecular approaches in regulatory monitoring programmes is still under development (see Chapter 5; Borja et al., 2024b).

[60] https://www.edito.eu/
[61] https://dnaqua.net
[62] https://northsearegion.eu/geans/index.html
[63] https://ednaquaplan.com/
[64] https://www.cost.eu/actions/CA25163/
[65] EOVs are a minimum set of Ocean variables that are needed to assess Ocean state and variability for important global Ocean phenomena (see GOOS phenomena list), and to provide essential data for applications that support societal benefit. https://goosocean.org/what-we-do/framework/essential-ocean-variables/
[66] EBVs are global standardised, scientifically-defined indicators developed to track key aspects of biodiversity and used to guide reporting on biodiversity within major treaties https://geobon.org/ebvs/what-are-ebvs/
[67] https://geobon.org/

For the WFD, molecular tools can support ecological quality classification by improving phytoplankton monitoring and detecting shifts in community composition associated with eutrophication and environmental stress. In the context of HABs, molecular methods are being integrated into environmental assessment and management frameworks under the WFD (see Box 2.6.) and the MSFD. Traditional monitoring approaches remain essential, but they are often time-consuming and can miss cryptic, toxic, or low-abundance species early on. Molecular approaches such as PCR-based methods (see Section 3.3.1.) provide rapid, sensitive, and species-specific detection of harmful microalgae and associated toxin-producing taxa, improving the identification of bloom dynamics, and enhancing the assessment of biodiversity and ecological status in coastal and transitional waters. qPCR methods for marine HAB species are widely used in research and operational HAB monitoring programmes worldwide. Implementation varies by region, depending on expertise, infrastructure, funding, regulatory frameworks, and bloom history. These methods are already adopted in Ireland, Scotland, France and Spain (John et al., 2026). qPCR results support decisions regarding biotoxin monitoring and have led to improvements in toxin management in shellfish aquaculture (Ruvindy et al., 2018). Within the WFD, metabarcoding has been tested for use in routine assessment of other biological elements, such as macroinvertebrates (Borja et al., 2024b) and fish (Bhendarkar et al., 2025), and some strategies have been proposed for successful integration of DNA-based methods into WFD monitoring (Blancher et al., 2022).

Box 2.6. The use of metabarcoding for monitoring and forecasting harmful algal blooms in support of the Water Framework Directive (WFD) and Marine Strategy Framework Directive (MSFD)



Credit: Ruđer Bošković Institute, Center for Marine Research

The toxic phytoplankton species, *Pseudo-nitzschia*.

The WFD and MSFD mandate EU Member States to monitor and manage the quality of inland, transitional and coastal waters. This may include tracking the presence and impact of HABs through phytoplankton monitoring in the WFD and is mandated in Descriptor 5 of the MSFD on eutrophication. Metabarcoding of algae in the shallow Ría de Vigo estuary of Northwest Spain, where intense aquaculture and fishery activities occur, enabled the simultaneous detection of 25 potentially harmful species from several algal groups included in the IOC-UNESCO Taxonomic Reference List of Harmful Microalgae[68], in both sediment and water samples, and across different seasons. Moreover, it showed that occurrence of HABs was influenced by upwelling and downwelling events (Ríos-Castro et al., 2023). This highlighted that including metabarcoding in monitoring systems allows tracking of seasonal changes in planktonic communities, providing early warning of potentially harmful species overlooked by conventional microscopical techniques. In addition, some toxic phytoplankton species, such as those within the genus *Pseudo-nitzschia*, cannot be reliably identified at the species level based on morphological features alone and may require molecular approaches for accurate identification.

[68] https://www.marinespecies.org/hab/

In addition, molecular methods have been successfully applied to improve biodiversity detection and ecosystem assessments under the EU Biodiversity Strategy by enhancing species inventories and enabling the detection of rare, cryptic and invasive species (Pawlowski et al., 2018; Aylagas et al., 2020). Similarly, eDNA and metabarcoding have been used in restoration projects to monitor changes in community composition and biodiversity recovery following habitat restoration, demonstrating their potential to support implementation of the EU Nature Restoration Regulation. Under the Urban Wastewater Treatment Directive (Directive 91/271/EEC) molecular approaches are being evaluated for the rapid detection of microbial contaminants and antimicrobial resistance genes. Likewise, qPCR and dPCR have been validated in pilot studies for the same-day detection of faecal indicator organisms and waterborne pathogens to support the Bathing Water Directive (see Box 2.7.). Under the CFP, eDNA and population genomic approaches have been applied in research and pilot monitoring programmes to complement conventional surveys by improving species detection, identifying spawning areas and supporting fish stock assessment (see Boxes 2.1. & 2.2.). Similarly, molecular methods have been widely used in conservation research to monitor protected species, assess habitat quality and detect invasive species, providing evidence that could support future implementation under the Habitats Directive.

Current research focuses on developing standardised protocols, harmonising biomolecular approaches through intercalibration exercises (see Section 5.2.), establishing reliable reference databases (see Section 3.3.3.), developing data-sharing infrastructures (see Section 5.6.) and quantitative frameworks (see Section 5.3.) that are crucial to scale-up biomolecular techniques for integration into routine monitoring programmes and policy implementation, while maintaining compatibility with long-term ecological datasets.

Box 2.7. The application of molecular tools to the Bathing Water Directive (BWD)



Credit: Lawrence Chismorie, Unsplash

The use of molecular methods can strengthen the monitoring of bathing water quality.

Under the European BWD Member States are required to routinely monitor, assess and publicly report bathing water quality, generating long-term datasets of microbiological and environmental parameters. These provide a valuable framework for comparing data collected using traditional methods with molecular approaches, particularly in coastal and transitional waters. In areas affected by pollution, sewage discharge, surface runoff or HABs, the integration of biomolecular methods such as qPCR in monitoring can provide more rapid assessment of faecal contamination by Escherichia coli and enterococci. This enables the communication of risk on the same day as the detection, supporting management decisions (Oliver et al., 2016). Molecular methods are complementary to official BWD workflows that use culture-based methods that require 24–48 hours (Ahmed et al., 2019). Biomolecular approaches are well suited for monitoring following heavy rainfall or sewage overflow events, and for providing supplementary evidence (WHO, 2021). Biomolecular methods offer the possibility of expanding routine monitoring beyond faecal indicator organisms to include selected waterborne pathogens, particularly viruses, where site-specific public health concerns justify additional surveillance (WHO, 2021). Although these approaches have not yet been routinely implemented under the BWD, they are being actively evaluated to support same-day water quality assessment and risk management (Gomez Cortes et al., 2024).

2.7. Recommendations

Recommendations for research funders:

(2.1) Fund co-designed research (involving policymakers, scientists and industry) linking biomolecular methods to policy applications, including co-development of new machine-readable policy indicators ready for integration into the European Digital Twin of the Ocean;

(2.2) Support the development of augmented observatories by including biomolecular data, alongside traditional observations, into longitudinal, meso-scale and process studies, allowing long-term calibration with traditional methods;

(2.3) Invest in easy-to-use metadata and data repositories, and visualisation interfaces for molecular data, to ensure broad use of scaled-up molecular observations, data sharing and accessibility in an open science framework;

(2.4) Fund research on the fundamental understanding of biodiversity and ecosystem function using biomolecular tools to stimulate potential advances in policy-relevant areas, and with potential practical applications in biotechnology;

(2.5) Fund projects on the novel interpretation of existing datasets to stimulate trans- and inter- disciplinary data analysis; and

(2.6) Fund proof-of-concept research on the use of biomolecular data to understand baseline ecological dynamics and the potential effects of industrial activities.

Recommendations for scientists:

(2.7) Establish a working group, potentially under the Group on Earth Observations Biodiversity Observation Network (GeoBON), on the cross-cutting implementation of biomolecular tools in all classes of Essential Biological Variables; and

(2.8) Develop a framework to value ecosystem services based on molecular functions.

Recommendations for policymakers:

(2.9) Integrate insights on biodiversity and ecosystem function from biomolecular data into the design of protected areas and restoration initiatives to support policies including the Habitats Directive and Nature Restoration Regulation; and

(2.10) Embrace novel frameworks to value ecosystem services based on molecular functions.

3 State of the art in biomolecular methods

3.1. Summary

This chapter outlines the state of the art of modern biomolecular methods from sample acquisition to the biological insights that the data provide, as well as their benefits and drawbacks (see Table 3.1. for overview). The chapter also discusses the use of biomolecular data in modelling and the importance of improving reference databases to accurately characterise the taxonomic origin and function of sequences.

3.2. Sources of molecular material

Molecular methods can target an organism or a community (see Figure 3.1.). Organismal samples include muscle tissue, fish fin clips, blood samples, and parts of algae, as well as entire small organisms like larvae, phytoplankton, and microscopic invertebrates. Samples can be obtained either through sampling in the field or from archived collections. Community samples from the field can be obtained by filtering water or extracting DNA from sediment samples. Nets (e.g. plankton nets) yield samples that contain organisms such as phytoplankton and zooplankton depending on mesh size, while filters with fine pores yield ciliates[69] and microalgae, and those with sub-micron pore sizes yield bacteria and archaea as well as environmental DNA (eDNA) from cells shed by larger organisms or from dead and degraded cells.

Isolating molecular material from eDNA samples can trace a wide range of marine organisms that are, or have been, present at the sampling location. Samples can contain gametes (i.e. reproductive cells), resting cysts (i.e. dormant structures), dead cells, tissue, faeces or other molecular signatures that drift with Ocean currents that are not necessarily indicative of a viable population at the sampling site. An understanding of the persistence of eDNA in the environment is therefore required to ensure that "presences" are correctly interpreted. Consideration must be given to DNA degradation kinetics, i.e. how eDNA changes over time as it is shed, breaks down, and transported through the water. Meta-analyses have shown that the interactions between DNA degradation and environmental conditions in the marine environment are complex (Jo & Minamoto, 2021) and further work is required to allow accurate assessments of biodiversity.

Community eDNA samples can be obtained from sediment via grabs or cores to obtain the molecular signatures of organisms living on the seafloor and the degraded molecules from organisms living in the water column that sink to the seafloor after they die. This mixture of molecules can, over time, be preserved in the sediment as sedimentary ancient DNA (aDNA) and provide information on past marine biodiversity (see Section 2.2.).

3.3. Methodological state of the art

A range of PCR-based and -omics methods are available, which can be applied to either a single organism or a community, and can be used to detect species presence (Who's there?; see Section 2.2.) and function (Who does what?; see Section 2.3.) (Figure 3.2.). The prefix "meta-" (as in metabarcoding, metagenomics etc.) refers to the analysis of molecules across entire communities of organisms in environmental samples, rather than from a single organism. Despite similarities in techniques, the scientific questions that can be answered by each method and its 'meta-' counterpart vary (Figures 2.1. and 3.2., Table 3.1.).

The basic building blocks of DNA are repeating nucleotides[70], each containing a nitrogenous base, which may be adenine (A), thymine (T), guanine (G), or cytosine (C). The arrangement of these bases into sequences of three is the genetic code that dictates, through RNA, the assembly of amino-acids into proteins and ultimately determines an organism's entire structure and function. DNA sequences can be read through a variety of methods. Early-developed sequencing techniques (e.g. Sanger sequencing[71]) are limited to relatively short DNA fragments (typically up to 1000 nucleotides), yet remain valuable when only one specific DNA sequence needs to be analysed, for example, to verify a genetic modification, identify a DNA fragment, or confirming a suspected genetic variant. The newer Next Generation Sequencing[72] (NGS) techniques are not limited in this way and can generate sufficient data to support whole genome sequencing, and are cheaper

[69] A ciliate is a single-celled eukaryotic microorganism covered in tiny, hair-like structures called cilia.
[70] A nucleotide is a molecule that serves as the fundamental building block of nucleic acids such as DNA and RNA.
[71] Sanger sequencing (also known as chain termination sequencing) is a DNA sequencing method that uses modified nucleotides to terminate DNA synthesis at specific points.
[72] Next Generation Sequencing refers to technologies that sequence millions of small fragments of DNA in parallel. Also referred to as 'massively parallel' sequencing. NGS tends to be cheaper and faster than Sanger sequencing but requires bioinformatic analyses to generate the end sequence.

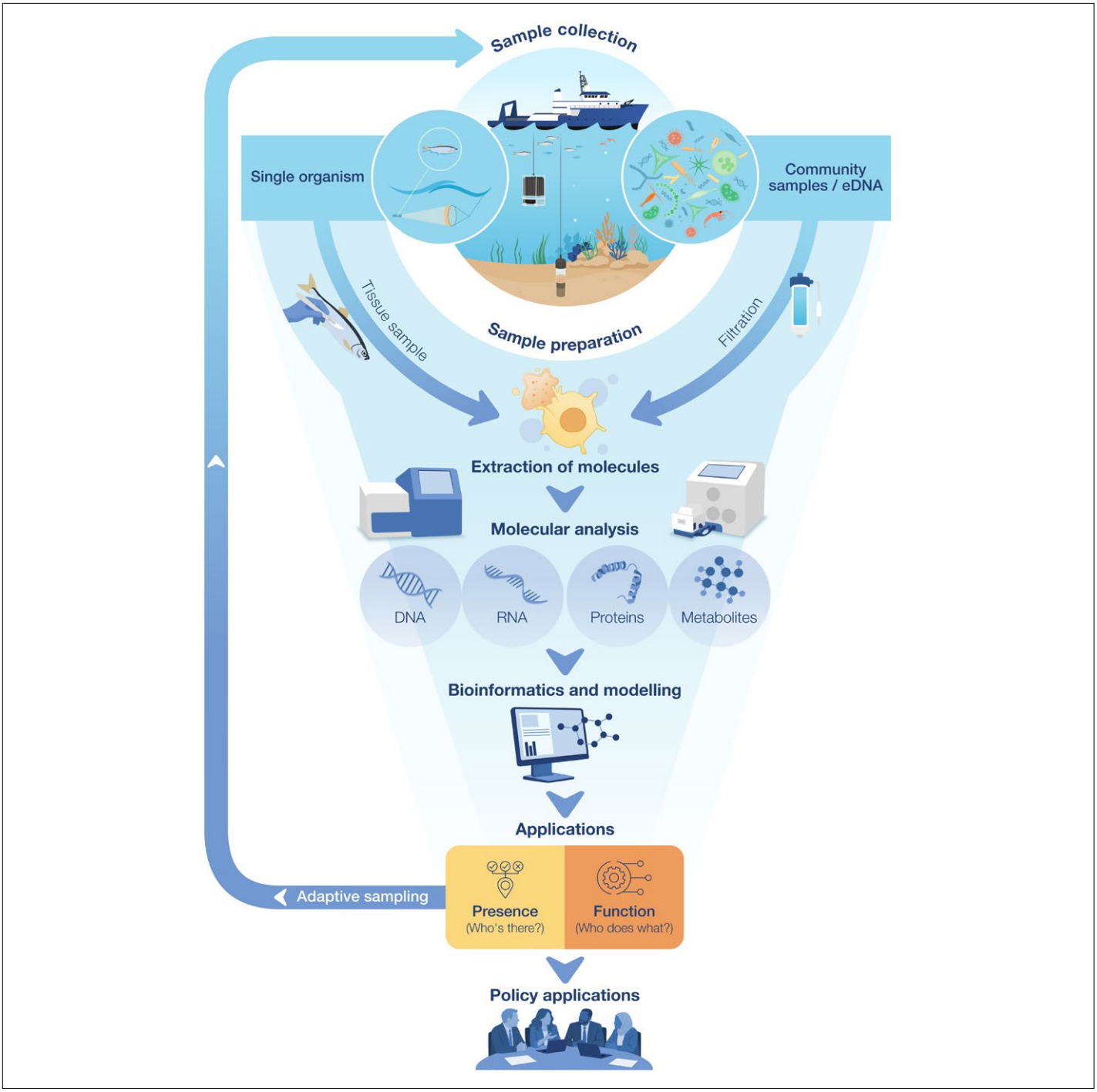


Figure 3.1. Overview of general workflow for biomolecular methods, from sampling to applications.

and faster than Sanger sequencing. The completeness and quality of the resulting genome sequence depend on factors including DNA quality, sequencing coverage, genome size and complexity, and the approach used to assemble the sequences. Short-read[73] NGS techniques (e.g. Illumina sequencing[74]) generate many short DNA sequences that can be assembled using bioinformatic approaches, either by comparison with a known reference genome or from scratch (*de novo*). Long-read[75] NGS platforms (e.g. PacBio[76], Oxford Nanopore[77]; see Section 4.3.) generate substantially longer sequences than conventional short-read platforms, which can simplify *de novo* assembly and resolve repetitive regions and structural variation. They do not eliminate the need for assembly, reference genomes or bioinformatic analysis, and their accuracy, yield and cost depend on the platform, chemistry and sample preparation.

[73] Short-read techniques are where a sample DNA is fragmented into short lengths prior to sequencing and reads of 50-300 base pairs are later bioinformatically assembled into longer sequences.

[74] https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology.html

[75] Long-read techniques involve sequencing outputs that are typically 10,000 – 100,000 base pairs.

[76] https://www.pacb.com/

[77] https://nanoporetech.com/

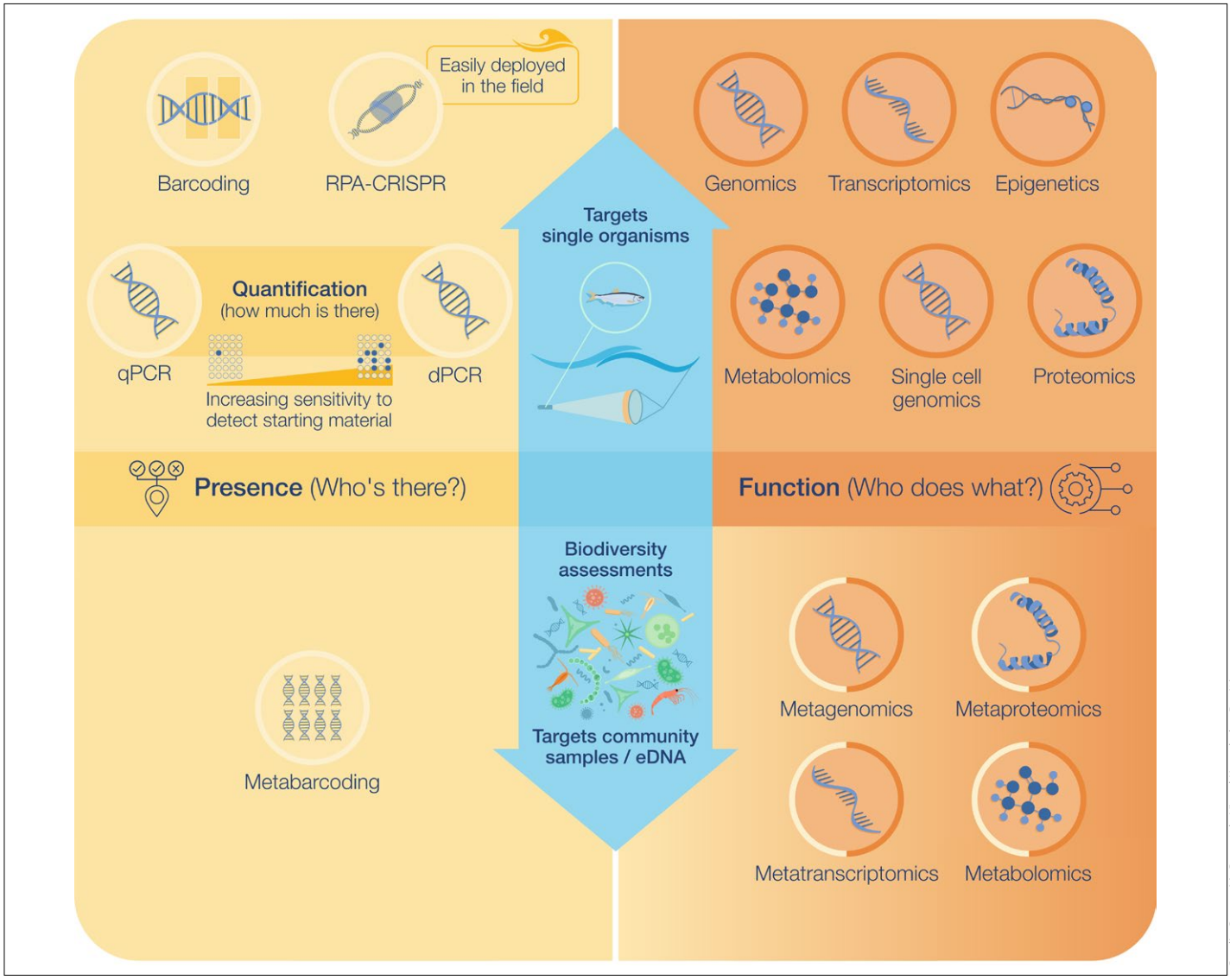


Figure 3.2. Molecular methods can be used to monitor and assess biodiversity, detect species presence, and determine functional responses of organisms and communities. Yellow colour coding indicates methods linked to determining the presence of organisms (i.e. Who's there?). Orange indicates methods linked to understanding organism and ecosystem function (i.e. Who does what?).

Polymerase Chain Reaction (PCR) methods produce copies of specific sections of DNA that can then be sequenced. PCR is a repeating temperature-controlled three-step process (Figure 3.3.). Heating denatures the doubled-stranded DNA into single strands (step 1), short strands of nucleotides known as 'primers' bind to matching target sequences on the single DNA strands (step 2), and finally an enzyme (Taq polymerase) and single nucleotides are added which generates a new DNA strand, resulting in new double-stranded DNA for the region between the primers (step 3). Repeating these steps leads to an exponential increase in the amount of target DNA, which is then easily detected or sequenced. If RNA is the starting material, incubation with an enzyme generates a matching DNA strand, followed by amplification using the three-step process described above. PCR can be used with Sanger sequencing when specific organisms are targeted, or with NGS methods when community samples are targeted.

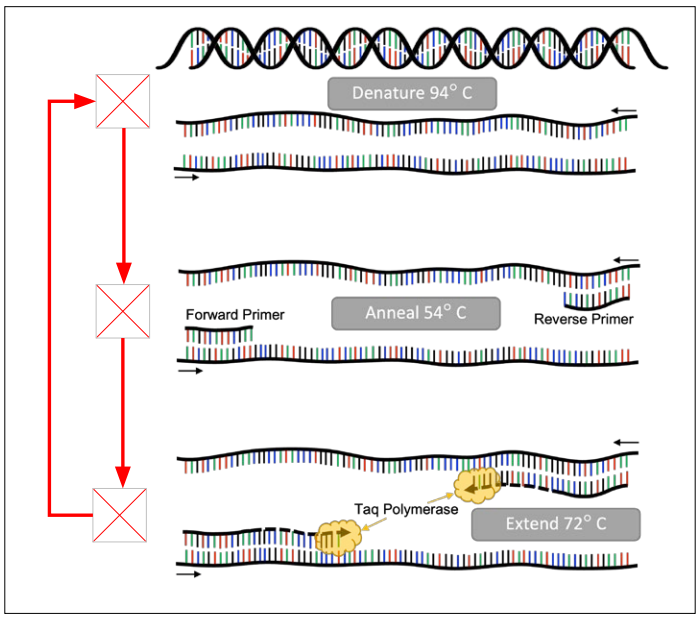
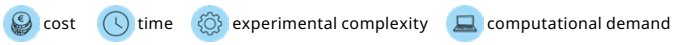


Figure 3.3. The repeating temperature-controlled three-step process of PCR to amplify specific DNA regions.

Table 3.1. Overview of biomolecular and -omics approaches used in environmental research, comparing approximate cost, time, experimental complexity, computational demand, applications, strengths and limitations. Cost and time depend on sample type, research question, number of sequences generated, instrumentation, and analytical workflows. Experimental complexity reflects the practical requirements for sample preparation, specialised equipment and laboratory procedures. "Field" and "onboard possible" indicate that the method can be adapted for deployment outside a conventional laboratory, however, this does not necessarily mean that all sample preparation and analysis can be completed in the field or onboard a ship. Computational demand reflects the requirements for data processing, analysis and interpretation. HAB = harmful algal bloom. NIS = non-indigenous species.



PCR/ QPCR			
€–€€	HOURS ONE DAY	LOW (FIELD/ ONBOARD POSSIBLE)	LOW
Applications Detection and quantification of HAB species and fish stocks; pathogen monitoring; detection of NIS; identification of antimicrobial resistance genes.	**Strengths** Fast; inexpensive; highly sensitive; targeted detection; absolute quantification; routine monitoring of species/genes.	**Limitations** Limited to predefined targets; no discovery of novel genes or taxa.	
dPCR			
€€	1-2 DAYS	MODERATE (LABORATORY/ ONBOARD POSSIBLE)	LOW-MODERATE
Applications Detection and absolute quantification of low abundance eDNA from rare species and organisms with low eDNA shedding; early detection of NIS, HABs and pathogens; quantification of microbes and viruses; monitoring pathogens and gene expression in aquaculture.	**Strengths** High sensitivity and precise quantification; suitable for low-abundance targets and reference measurements.	**Limitations** Targeted approach; lower throughput than sequencing; requires specialised instrumentation.	
ddPCR			
€€	1-3 DAYS	MODERATE-HIGH (LABORATORY)	LOW-MODERATE
Applications High-sensitivity quantification of rare DNA/ RNA targets; validation of eDNA surveys; monitoring viral loads and antimicrobial resistance markers.	**Strengths** High sensitivity and precise quantification; suitable for low-abundance targets and reference measurements.	**Limitations** Targeted approach; lower throughput than sequencing; requires specialised instrumentation.	
RPA-CRISPR			
€–€€	<1 HOUR	LOW (FIELD-DEPLOYABLE)	LOW
Applications Rapid field detection of pathogens; early warning of NIS; aquaculture disease surveillance; portable environmental diagnostics.	**Strengths** Rapid; highly sensitive and specific; portable; well suited for field and point-of-care applications.	**Limitations** Limited to predefined targets; less standardised than PCR; not suitable for discovery-based.	
METABARCODING			
€€	DAYS-WEEKS	LOW–MODERATE (LABORATORY)	LOW–MODERATE
Applications Biodiversity assessments; microbiome characterisation; HAB monitoring; community profiling of fish, plankton, benthos, fungi and other eukaryotes; reconstruction of past biodiversity using ancient eDNA; NIS detection.	**Strengths** Community composition; relatively inexpensive; standardised workflows; high throughput.	**Limitations** Primarily relative abundances; limited taxonomic resolution for some groups; little functional information; PCR amplification bias can distort estimates of community composition.	

(META)GENOMICS

€€€	WEEKS	MODERATE (LABORATORY)	HIGH

Applications
Characterisation of microbial community structure and functional potential; novel gene discovery; antimicrobial resistance surveillance; reconstruction of metagenome-assembled genomes; ecosystem monitoring.

Strengths
Comprehensive taxonomic and functional potential; detection of novel genes and organisms; strain-level resolution possible; can estimate absolute abundances.

Limitations
Functional potential rather than activity; absolute quantification requires internal standards.

(META)TRANSCRIPTOMICS

€€€–€€€€	WEEKS	HIGH (LABORATORY)	VERY HIGH

Applications
Identification of active microbial functions; ecosystem responses to environmental change; detection of RNA viruses; characterisation of active metabolic pathways.

Strengths
Active gene expression and metabolic pathways; detects RNA viruses; high functional resolution; can estimate absolute transcript abundances.

Limitations
RNA instability; technically demanding; reflects a snapshot in time; transcript abundance does not necessarily correspond to protein activity.

(META)PROTEOMICS

€€€€	WEEKS-MONTHS	VERY HIGH (SPECIALISED LABORATORY)	VERY HIGH

Applications
Characterisation of enzyme activity and expressed metabolic pathways; assessment of ecosystem functioning; biomarkers of environmental stress; host-microbiome interactions.

Strengths
Direct measurement of produced proteins and enzymes; closest to functional activity.

Limitations
Difficult protein extraction and identification; incomplete reference databases; lower proteome coverage; expensive instrumentation.

METABOLOMICS

€€€–€€€€	DAYS-WEEKS	HIGH (LABORATORY)	HIGH

Applications
Identification of biomarkers of ecosystem health; characterisation of carbon and nutrient cycling; organismal stress response; chemical ecology; discovery of natural products.

Strengths
Direct measure of biochemistry and metabolic state; detects rapid physiological responses; compliments other -omics.

Limitations
Challenging annotation; many metabolites remain unidentified; metabolite concentrations change rapidly.

EPIGENETICS

€€€€	WEEKS	HIGH–VERY HIGH (LABORATORY)	HIGH–VERY HIGH

Applications
Environmental adaptation and acclimation; stress response; transgenerational effects; population resilience to climate change; age determination in fish.

Strengths
Regulatory mechanisms and environmental adaptation without DNA sequence changes; genotype-phenotype link.

Limitations
Poorly standardised methods; complex interpretation; ecological significance often difficult to infer.

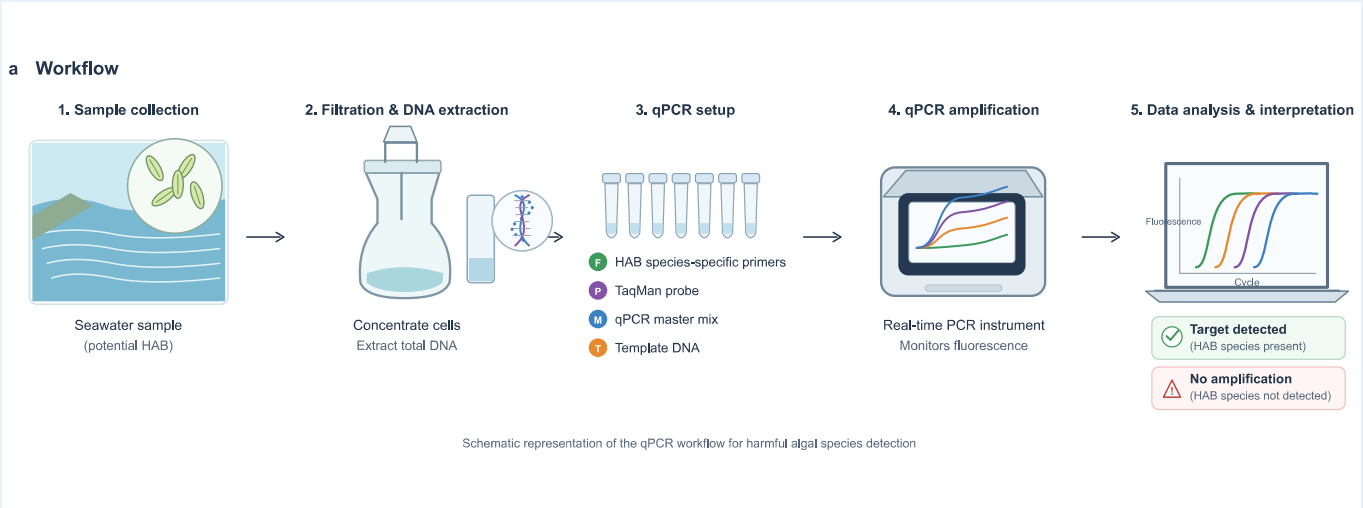
3.3.1. PCR-based methods

3.3.1.1. Quantitative PCR

While PCR produces copies of specific sections of DNA, quantitative PCR (qPCR) allows the amount of these specific sections of DNA to be quantified in real-time. This allows quantification of target species (e.g. viral and microbial loads) or gene expression levels in environmental samples. It has extensive applications including for harmful algal bloom (HAB) detection (see Box 3.1.), pathogen detection, eDNA-based biodiversity assessments, and the analysis of gene expression of organisms under environmental stress.

The technique involves the amplification of DNA using species- or gene- specific primers and short pieces of DNA tagged with a fluorescent dye (e.g. TaqMan probe), generating a fluorescence response that is directly proportional to the amount of DNA copies, allowing DNA quantification. The results are compared to standard curves, i.e. charts made using DNA from reference genes at known concentrations versus fluorescence signal, to calculate the amount of DNA in the test sample. qPCR is highly sensitive, specific, reproducible, and produces data rapidly and at low cost (Table 3.1.). Reliable qPCR depends on careful experimental design, including the selection of suitable controls for calibration, the optimisation of primer and probe sequences, and ensuring that the target DNA is amplified accurately and consistently. It is therefore normally limited to lab-based applications and selected field applications.

Box 3.1. qPCR for the early detection of harmful algal blooms (HABs)



General workflow for qPCR detection of HAB species.

Molecular monitoring is generally used as a lab-based rapid diagnostic tool for HABs. Water samples are collected according to surveillance programmes, typically weekly during periods of elevated HAB risk, with more frequent sampling during bloom events. Samples are transported to laboratories where DNA is extracted and analysed by qPCR, allowing species-specific results to be obtained within a few hours (see figure). This rapid turnaround enables same-day management decisions, such as increasing surveillance, initiating toxin analyses, implementing precautionary harvesting (e.g. in aquaculture) or bathing restrictions. Consequently, qPCR is increasingly viewed as a method for rapid early-warning and a decision-support tool that complements traditional HAB monitoring that uses light optical microscopy.

Within the EU official-controls framework for shellfish safety, competent authorities are required to monitor toxin-producing plankton in areas where shellfish are grown (Commission Implementing Regulation (EU) 2019/627). While this framework mandates the monitoring itself, it does not prescribe molecular methods. In this context, qPCR has increasingly moved from research towards complementary operational applications alongside established monitoring approaches, although it is not yet routinely implemented under EU legislation and its degree of uptake varies among countries and monitoring programmes (John et al., 2026).

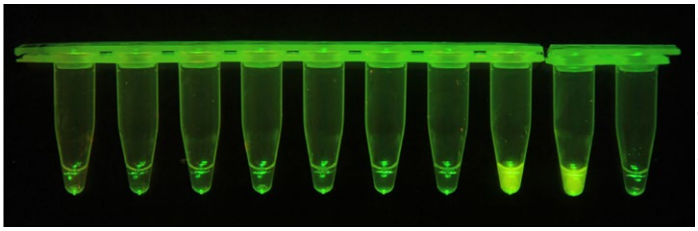
3.3.1.2. Digital PCR and droplet digital PCR

Digital PCR (dPCR) is the next step up from qPCR in terms of accuracy and complexity since it is able to detect signals that may be missed by qPCR. It allows absolute DNA quantification (rather than relative quantification against standard curves) and ultra-high sensitivity, which is useful to detect extremely dilute DNA, and when samples contain other molecules that can interfere with the reaction (Table 3.1.). It is ideal for detecting low quantities of eDNA from rare, elusive or cryptic species, and species with low eDNA shedding such as marine mammals that reduce their rate of skin shedding in response to cold temperature (Pinfield et al., 2019). It can also be used for early detection of invasive species, HABs, pathogens (see Box 2.5.), for quantifying microbes and viruses, and to detect slight increases in pathogen load and gene expression in aquaculture.

dPCR methods partition the sample into thousands of individual PCR reactions within a microfluidic system (droplet dPCR; ddPCR) or on a microfluidic chip with nanoscale chambers (cdPCR). During PCR amplification, droplets or chambers that contain copies of the target DNA sequence fluoresce. Statistics are applied to the data on the amount of fluorescence to estimate target molecule concentration so that no standard curve is needed. Because dPCR is highly sensitive, positive results must be carefully interpreted in the context of relevant ecology and taxonomy. Presence of DNA in the environment does not necessarily mean the presence of viable organisms at the time of sampling since DNA can persist after the death of organisms and can be transported from its sources by water masses and currents.

3.3.1.3. RPA-CRISPR

RPA-CRISPR is a method to detect species-specific presence or abundance. It is used to detect toxin-producing plankton in shellfish harvesting areas, to monitor levels of human pathogens in coastal waters in support of the Bathing Water Directive (BWD), and in the early and rapid detection of non-indigenous species in the field (see Box 3.2.). RPA-CRISPR is a two-step process. First, species-specific primers are used to amplify DNA from the target organism using recombinase polymerase amplification (RPA), which operates at a constant, relatively low temperature of 37-42°C. Then a CRISPR-Cas system detects a specific DNA sequence within the amplified DNA and quantifies it based on fluorescent signals. RPA-CRISPR can be used to accurately and rapidly detect species in the field within 10-20 minutes. It is easy to use and interpret results using fluorescence or lateral flow strips (i.e. paper-based devices used to detect presence of a substance). However, it requires a unique test for each target species, which necessitates expertise and resources.



Credit: Adapted from Zhao et al. (2024) (CC BY 4.0).

RPA-CRISPR is used to detect species and their abundance through specific DNA sequences that are quantified based on fluorescent signals.

Box 3.2. RPA-CRISPR for the early and rapid field based detection of non-indigenous jellyfish



Credit: Flickr (CC BY-2.0)

Japanese sea nettle *(Chrysaora pacifica).*

The non-indigenous Japanese sea nettle Chrysaora pacifica can multiply quickly and cause extensive economic damage by reducing the food available for developing fish larvae and killing fish in aquaculture enclosures in South Korea. Kim et al. (2024) used eDNA and RPA-CRISPR to detect the presence of Chrysaora pacifica, which correlated with results from dPCR analysis showing the potential for rapid, on-site early detection of harmful jellyfish blooms. If developed further into user-friendly commercial formats, it could be deployed by trained personnel involved in coastal and marine monitoring. This method could be adapted for problematic European jellyfish species such as the mauve stinger *Pelagia noctiluca*, to provide early warning for North Atlantic salmon farms, which are impacted by jellyfish blooms (Bosch-Belmar et al., 2021). Early detection allows fish farms to deploy bubble curtains or fine-mesh nets to provide physical protection, or to make decisions on early harvesting.

3.3.1.4. (Meta)barcoding

DNA barcoding is a technique to amplify a fragment of an organism's genome, known as a barcode, that varies between species but is similar within a species. It can be used to reliably distinguish and identify species for biodiversity assessments, and detect community shifts and ecological interactions (see Section 2.2.). It can also be applied to marine protection plans and non-indigenous species detection (see Box 3.3.). Metabarcoding can be used to analyse community samples and can be applied to eDNA. It enables taxonomic identification of the species within a community through amplification of a selected DNA barcode. The choice of marker region and primers for amplification depends on the taxonomic scope and target groups of study, as different markers and primer sets vary in their capacity to detect certain taxonomic groups, amplification efficiency, and discriminatory power (see Table 3.2.). The selected barcode region is amplified by PCR, producing DNA fragments that are subsequently sequenced using next-generation sequencing, followed by taxonomic assignment via bioinformatic comparison of the sequences to curated reference databases specialised for specific barcoding genes (see Section 3.3.3. and Table 3.2.). As databases of barcodes for particular organisms expand, these can be used for other purposes. For example, Fishbase has released a Genetic Diversity Risk Indicator tool[78] for commercial fish populations.

Barcoding a single organism is simpler than metabarcoding complex communities because the PCR typically produces a single dominant DNA sequence, requiring little downstream bioinformatic processing. In contrast, metabarcoding of community samples generates many DNA sequences that require more extensive bioinformatic processing but enables rapid, reliable, and comprehensive taxonomic characterisation of complex communities. High-throughput and cost-effective sequencing technologies provide low-cost and rapid assessments. Rare taxa, species present in low abundances, and taxa/species that are difficult to identify with traditional morphological methods can be successfully identified.

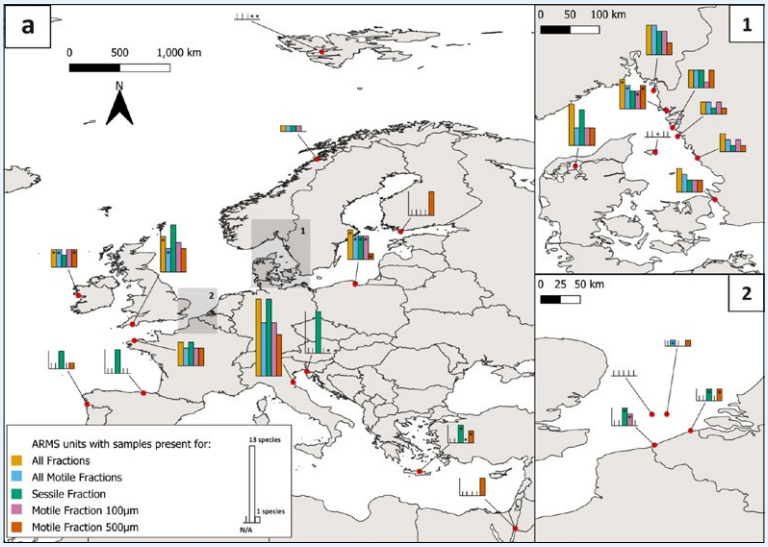
It is essential to follow standardised protocols to minimise variation, ensure consistency across samples and allow comparison of results between different studies (see Section 5.2). In addition, metabarcoding is only semiquantitative, providing information only on relative abundances. However, studies indicate that the relative abundances of DNA obtained through metabarcoding can serve as a proxy for the relative abundances of taxa in the studied environment (Rourke et al., 2022).

Table 3.2. The most commonly used barcoding genes with their corresponding curated reference databases, and the target group of organisms for each gene and database.

	TARGET GROUP	GENES/BARCODING REGIONS (Abbreviations refer to names of genes/barcoding regions; full descriptions are in the Abbreviation list)	DATABASE
Eukaryotes	Animals	COI	BOLD[79]
	Fish	COI, 12S	Mare-MAGE[80] MitoFish[81]
	Zooplankton	COI, 12S, 16S, 18S, 28S	MZGdb[82]
	Fungi	ITS	UNITE[83]
	Diatoms	rbcL	Diat.barcode[84]
	Protists	18S	PR2[85]
	Metazoa, Protists	18S, 28S	SILVA[86]
	Photosynthetic protists	Plastidial 16S	Phytoref[87]
Prokaryotes[88]	Bacteria, Archaea	16S	SILVA
	Bacteria, Archaea	16S	GreenGenes2[89]
	Bacteria, Archaea	16S	GTDB[90]

78 https://www.fishbase.ca/Genetics/GeneticDiversityTool.php
79 https://boldsystems.org/
80 https://mare-mage.weebly.com/
81 https://mitofish.aori.u-tokyo.ac.jp/
82 https://metazoogene.org/mzgdb/
83 https://unite.ut.ee/
84 https://carrtel-collection.hub.inrae.fr/barcoding-databases/diat.barcode
85 https://pr2-database.org/
86 https://www.arb-silva.de/
87 http://phytoref.sb-roscoff.fr/
88 Prokaryote are unicellular organisms (i.e. Bacteria and Archaea) lacking a membrane-bound nucleus and membrane-bound organelles.
89 https://github.com/biocore/greengenes2
90 https://gtdb.ecogenomic.org

Box 3.3. Case study on the use of metabarcoding to detect non-indigenous species (NIS)



Credit: Pagnier et al. (2025) (CC BY 4.0).

Comparative distribution of NIS detected using metabarcoding across the European Autonomous Reef Monitoring Structures (ARMS) Marine Biodiversity Observation Network. Bar charts show the number of NIS per site across five fractions of organisms. Insets show areas with multiple sampling sites.

The implementation of metabarcoding as a tool to detect newly introduced NIS and the geographic spread of established NIS could lead to earlier detections and more effective responses. For example, the Autonomous Reef Monitoring Structures (ARMS) which is part of the Marine Biodiversity Observation Network (MBON) (ARMS-MBON) are stacks of settlement plates attached close to the seafloor and metabarcoding is used to detect marine organisms that colonise the plates. 112 of these structures were installed across 19 benthic observatories in 14 countries across Europe. 1,828 species were identified by comparing DNA sequences against established reference databases (Pagnier et al., 2025). Further comparison of the identified species against the World Register of Introduced Marine Species[91] database established that 63 of the species were NIS. NIS were present on almost every plate (see figure). Individual NIS species were often widespread, and detections at some locations indicated further spreading of particular NIS species. Such detections include the first record of bryozoans *Bugula neritina* in the Red Sea, and *Fenestrulina delicia* in the Mediterranean. Despite the high number of detections, a very large number of sequences could not be identified due to gaps in reference databases, highlighting the urgent need for their improvement (see Section 3.3.3.). Comparative studies are currently evaluating species identification rates in combination with traditional taxonomic identification (Varrella et al., 2025).

3.3.2. -Omics approaches

-Omics approaches (e.g. genomics, transcriptomics, proteomics) are used to study entire sets of biological molecules (e.g. DNA, RNA, proteins) to understand how organisms function. In practice, environmental -omics is mainly microbial, because microbial genes and proteins dominate many environmental samples and collectively, microorganisms comprise the second largest living biomass on Earth, after plants (Bar-On et al., 2018), and are the engines driving biogeochemical cycles. However, mixed-community -omics, and -omics of host-associated microbiomes (e.g. for coral, sponges, mussels) allow detection of host and microbial molecules, making it useful beyond microorganisms alone.

-Omics methods depend on the availability and quality of reference databases (see Section 3.3.3.), which are better developed for microbes than other organisms, and better developed for some biomolecular methods than others. Metagenomics analyses eDNA to determine what could happen, metatranscriptomics analyses environmental RNA to determine what genes are being expressed, and metaproteomics analyses environmental proteins to determine active functions of an organism. Together, these provide complementary insights into ecosystem structure and function and their integration into multi-omics approaches can enhance understanding of ecosystem function and community interactions (see Section 4.4.).

3.3.2.1. (Meta)genomics

Metagenomics analyses the total DNA present in a community sample (e.g. from eDNA) without targeted PCR amplification. This enables the characterisation of the composition and potential functional capabilities of biological communities. Importantly, many environmental microorganisms cannot readily be cultivated in laboratories, thus their biology cannot be studied using traditional approaches. Metagenomics can enable the reconstruction of complete or near-complete genomes of uncultivable microorganisms (Albertsen et al., 2013). The reconstruction of genomes allows genes and metabolic pathways to be associated with individual microbial populations. This enables investigation of microbial metabolic capabilities and interactions, as well as exploration of genes with potential biotechnological applications (He et al., 2021). Metagenomics also enables the detection of species that might be missed by metabarcoding (Callens et al., 2025). The resulting sequences are analysed bioinformatically to identify and annotate genes, allowing potential functions and metabolic pathways to be inferred.

Metagenomics has become increasingly accessible owing to advances in high-throughput sequencing and standardised laboratory workflows, enabling rapid and comprehensive characterisation of microbial communities and their functional potential. Metagenomics is more costly than metabarcoding (Table 3.1.) because it generates larger and more complex

[91] https://www.marinespecies.org/introduced/

datasets that require substantial computational resources and bioinformatic expertise (Laudadio et al., 2019). Metagenomics determines functional potential rather than actual activity or ecological relevance, emphasising the complementary value of metatranscriptomics (Section 3.3.2.2.) and other multi-omics approaches (see Section 4.5.).

Genomics (as opposed to metagenomics) focuses on individual organisms and is widely applied across diverse taxa, including plants and animals, whose genomes are often large and structurally complex. Genomics provides high-resolution insights into evolution, reliable information on gene function, and data to populate reference databases. The choice between genomics and metagenomics therefore depends on the organisms and research questions of interest: genomics is suited to investigating individual organisms in detail, whereas metagenomics enables the study of entire communities.

Flow sorting instrumentation that allows the sorting of single cells has been adopted in marine science to sequence genomes of targeted taxa and to understand interactions. Individual cells are sorted directly from environmental samples, such as seawater, before their genomes are amplified and sequenced. This enables targeted genomic information to be obtained from uncultured rare organisms with important ecological roles that may be difficult to characterise accurately through large-scale sequencing of bulk environmental samples.

Several interesting questions can be answered with genomics without the need to sequence and assemble an entire genome. Ultra conserved elements[92] and Single Nucleotide Polymorphisms[93] (SNPs) are easily extracted from partial genome sequences and can be analysed to answer questions on defining species, genetic diversity, and gene flow among populations. For example, estimates of effective population size of European Hake, *Merluccius merluccius*, derived from SNPs suggest drastic declines in genetic diversity over the last 50 years (Petit-Marty et al., 2025). Similarly SNPs have revealed extremely low genetic diversity in Black Sea harbour porpoises, *Phocoena phocoena*, and identified the critically endangered Baltic Sea harbour porpoise as a genetically isolated population (Celemín et al., 2025). Such studies can act as early-warning signals and support appropriate management measures to maintain genetic diversity and adaptive potential in exploited species managed under the CFP and species of conservation concern under the Habitats Directive.

3.3.2.2. (Meta)transcriptomics

Transcriptomics identifies genes that are actively expressed and enables high-resolution comparisons of gene expression among tissues, across developmental stages or time points, and in response to changing environmental conditions. Transcriptomic analyses can be carried out on individual organisms or can be applied across a broad range of taxa.

Metatranscriptomics, also referred to as environmental transcriptomics, enables monitoring of gene expression across virtually all genes in a sample without prior selection of target genes. This is particularly useful for the direct assessment of functional responses of biological communities across diverse environmental systems (Anwar et al., 2019). Metatranscriptomics provides functional "snapshots" of physiological activity by measuring gene expression from samples taken at a particular location and time. It can be used to identify active genes, detect organisms contributing to ecosystem functioning, and characterise community-level responses to environmental change.



Credit: Jorge Juan Santos Bruña.

Metatranscriptomics provides functional "snapshots" of physiological activity by measuring gene expression and characterising community responses to environmental change.

[92] Ultra Conserved Elements are segments of DNA that are identical or nearly identical across different species.

[93] A Single Nucleotide Polymorphism (SNP) is a single-letter difference in the DNA sequence among individuals or populations. SNPs are commonly used as genetic markers.

RNA sequencing (RNA-seq) is one of the most used methods for metatranscriptomics. In a typical workflow, RNA molecules (e.g. messenger RNA, mRNA) are reverse-transcribed into complementary DNA (cDNA), which is then sequenced and analysed to identify expressed genes and infer their potential functions and taxonomic identities.

Metatranscriptomics is the only approach that can detect entities whose genetic material consists exclusively of RNA, such as RNA viruses and viroids[94] making it indispensable for exploring RNA-based biodiversity (Koonin & Lee, 2025). In contrast to targeted approaches such as dPCR, metatranscriptomics does not require primers or probes for specific genes and therefore provides orders of magnitude more information on the functional activity and regulation of microbial communities.

However, metatranscriptomics is more resource-intensive than metagenomics in terms of laboratory processing and sequencing costs (Table 3.1.) and RNA is chemically unstable and highly susceptible to degradation, which can introduce biases (Kornienko et al., 2024). In addition, mRNA abundance does not always directly reflect protein levels or enzymatic activity (Moran et al., 2013). Reliable metatranscriptomic measurements therefore require careful consideration of when and where samples are collected, how samples are preserved and processed, and the use of standardised protocols to minimise technical variation. Integrated multi-omics approaches can further help to link gene expression to protein production (see Section 4.4.).

Laboratory- and aquarium-based studies are essential for the identification and characterisation of genes of unknown function, and to understand the response (typically gene expression) of organisms to stressors under controlled conditions. While controlled lab conditions are different from the environment, characterising molecular responses to discrete manipulations are only possible in the lab, and are required to link transcriptional patterns to their underlying causes. These transcripts or transcriptional patterns are measured *in situ* to diagnose responses to specific stressors. This type of research has accelerated our understanding of microbiomes (Hutchins & Fu, 2017), which is required to advance environmental -omics monitoring. Continued lab research from microbiomes to metazoans is necessary to expand the knowledge gained from the vast number of samples that will be collected as autonomous platforms reach their scaling potential (see Section 4.2.).

3.3.2.3. (Meta)proteomics

Proteomics is used to characterise and quantify proteins, providing information on the functional machinery present in organisms and systems at a given moment. In contrast to metatranscriptomics, which measures gene transcripts and therefore indicates which genes are being actively transcribed, proteomics measures the proteins produced from those genes and can provide a closer link to cellular function and physiological and metabolic states. Changes in protein composition over time can therefore reflect responses to environmental conditions and regulatory processes. This capability is particularly important for understanding complex and rapidly changing biological processes.

Metaproteomics extends these capabilities to complex biological communities, enabling simultaneous measurements of functional activity across multiple coexisting organisms in environmental samples (see Box 3.4.). For example, by directly observing expressed metabolic pathways and interactions, metaproteomics is particularly useful in providing system-level insights into microbial cooperation, resource use, and adaptive responses. It has been used to identify functional indicators of ecosystem responses to environmental change, including oxygen depletion, pollution, and climate-related stress by identifying changes in expressed proteins and metabolic pathways (see Box 3.4.; Tomanek, 2011). These functional indicators can complement biodiversity assessments by providing information on ecosystem functioning and resilience. Beyond environmental applications, proteomics can also support fisheries and aquaculture management by assessing physiological condition, health, nutrition, and environmental stress responses (Rodrigues et al., 2018).

Proteomic analyses typically involve protein extraction, enzymatic digestion, and analysis using high-resolution mass spectrometry to measure the mass of molecules and identify their chemical structure (Yates et al., 2009). Protein identification and quantification are achieved by matching mass spectra to reference databases, e.g. UniProt[95], using computational algorithms, enabling the reconstruction of functional protein profiles in biological and environmental samples.

Proteomics also supports rapid species identification. For example, matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) generates characteristic protein mass spectra that function as molecular fingerprints. This enables fast, cost-effective identification of species, including morphologically similar or cryptic taxa that



Credit: Pixabay.

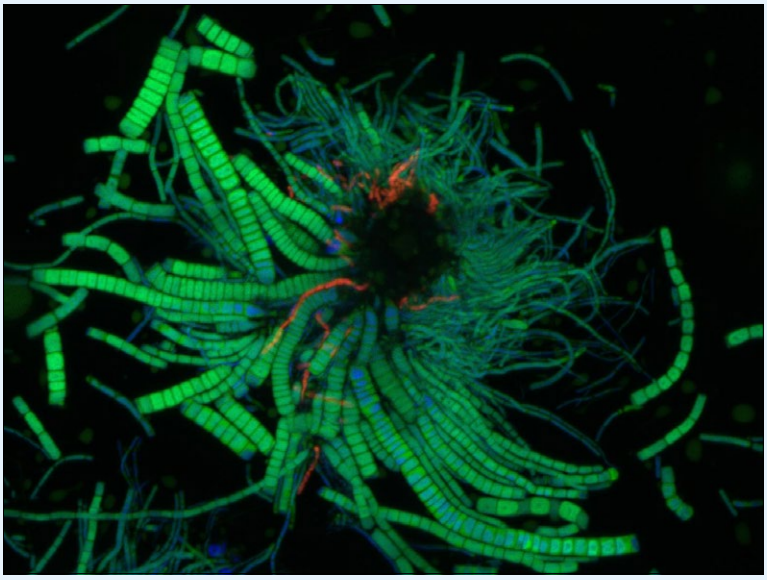
Proteomics can also support fisheries and aquaculture management by assessing physiological condition, health, nutrition, and environmental stress responses.

[94] A viroid is a small, infectious agent made up of a short strand of RNA with no protein shell.

[95] https://www.uniprot.org/

are difficult to distinguish using visual or genetic methods alone. However, marine proteomic analyses face specific challenges, mainly related to high salt concentrations in samples, which can affect protein stability and mass spectrometry ionisation efficiency (Cañas et al., 2007). Therefore, proteomic workflows are technically more complex, costly, and sensitive to sample handling and degradation than genomic workflows (Table 3.1.). Proteomics and metaproteomics capture only a subset of expressed proteins, compared with complete genomes or transcriptomes that code for all, and all expressed proteins respectively. In particular, low-abundance, hydrophobic, and basic proteins are not well covered. In addition, protein abundance often reflects steady-state conditions and may lag behind rapid physiological changes captured more easily in RNA.

Box 3.4. Metaproteomics to investigate how the Ocean microbiome processes organic matter



Credit: Ifremer (CC BY 4.0).

Metaproteomics can be used to understand how marine microorganisms process organic matter to predict Ocean carbon cycling.

Understanding how marine microorganisms process organic matter is essential for predicting Ocean carbon cycling and ecosystem functioning. Metaproteomics can be used to directly observe the metabolic activities occurring in natural marine environments e.g. through the identification of thousands of expressed proteins in microbial communities collected from depths ranging from 100 to 5,000 metres in the Atlantic Ocean (Bergauer et al., 2018). This approach provided insights into the importance of organic matter for deep-sea microorganisms by showing the high prevalence of proteins involved in the uptake of dissolved organic matter in the deep sea. Understanding how these microorganisms acquire and process organic matter is important because they play a major role in Ocean biogeochemical cycles and ecosystem functioning. By linking microbial identity to expressed metabolic functions, metaproteomics reveals the mechanisms underlying Ocean biogeochemical cycles and enables direct observation of microbial ecosystem functions, advancing our ability to understand and monitor the biological processes that regulate carbon flow and nutrient cycling in the Ocean. Such information could support MSFD assessments, particularly D5: eutrophication, by improving our understanding of microbial processes involved in nutrient and organic matter cycling, and D4: food webs, by providing insights into the functional roles of microorganisms within marine food webs.

3.3.2.4. Epigenetics / Epigenomics

Epigenetics refers to heritable changes in gene expression and cell function that occur without altering the DNA sequence, controlled through chemical modifications that act as molecular switches, activating or deactivating genes. These include adding chemical tags known as methyl groups (i.e. methylation) on the DNA molecule; the modification of histones (i.e. proteins in the nucleus) through processes such as acetylation (i.e. addition of acetyl group) and methylation; and the regulation of gene expression through non-coding RNA molecules (i.e. that do not code for proteins and instead have regulatory functions).

The expression patterns of epigenetic markers change in response to marine stressors such as acidification, deoxygenation, temperature fluctuations, and exposure to pollutants, making them candidates for early-warning biomarkers for environmental monitoring. For example, specific non-coding RNAs have been identified that regulate reproductive function and stress responses to oxygen depletion in a laboratory model fish (Lai et al., 2016), while changes in methylation have been used to identify the age of other fish species (see Box 3.5.). Epigenetic changes can be quantified from non-destructive tissue samples such as fish fin clips or blood and/or serum samples affording possibilities for non-destructive sampling. These biomarkers can potentially be used to assess ecosystem health, track the effects of climate change on commercially important species, inform fisheries management, and evaluate the effectiveness of conservation interventions before they become apparent at the population level. However, the development of epigenetics for policy is in its infancy and extensive further research and development is required to develop epigenetic techniques into a policy-relevant toolbox.

Box 3.5. Epigenetic clock analysis for fisheries management



Credit: Ivan Hurzeler and DEEP SEARCH 2019 BOEM, USGS, NOAA., ROV Jason ©Woods Hole Oceanographic Institution

DNA methylation in some fish species (e.g. deep-sea black-bellied rosefish) correlates with chronological age.

The age of individual fish is important for fisheries management, but it is not always easy to estimate. Scientists generally count growth marks on hard body parts, such as otoliths, scales, vertebrae or fin rays, which can be difficult to interpret in some species. Estimating age is particularly difficult in deep-sea fish, as they live for many years, grow slowly, and exist under more consistent environmental conditions that may provide a less pronounced annual signal. Epigenetic clocks are based on DNA methylation, an epigenetic mark that can change during an organism's lifetime. Studies have shown that methylation at certain sites in some fish species correlates with chronological age. For example, this was shown for the long-lived deep-sea black-bellied rosefish, Helicolenus dactylopterus, by Weber et al., (2024). Methylated sites were identified in both fin and muscle tissues in which DNA methylation changes with age. The methylation patterns were used to predict the fish's age and were compared with age estimates generated by counting the opaque zones in thin sections of otoliths. Ages predicted from DNA methylation were very close to the ages generated through otolith ring counts, especially when DNA was extracted from fin samples. This research demonstrates that epigenetic tools are useful for determining age in fish where traditional age determination is difficult and could provide useful age information for stock assessment under the CFP.

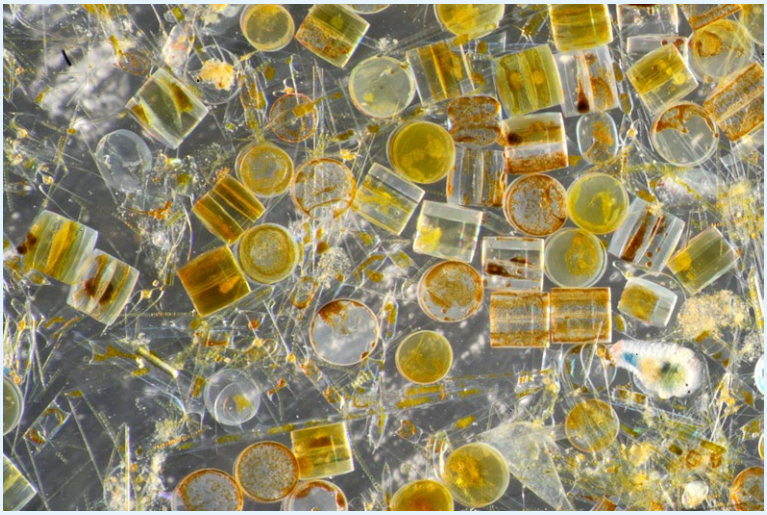
3.3.2.5. Metabolomics

The metabolome is the complete set of small molecules within an organism, including both internally produced compounds and those acquired from the environment (Patti et al., 2012). Unlike the genome, transcriptome, or proteome, which describe biological potential and machinery, the metabolome reflects biochemical state and ongoing metabolic processes[96]. Metabolites are broadly classified into primary metabolites, which are essential for growth and maintenance, and secondary metabolites, which mediate interactions such as signalling, defence, and mutually beneficial relationships between organisms. These include diverse specialised metabolites and lipid-derived compounds that can mediate interactions between organisms. Lipid composition can provide information on structural, energetic, signalling, and physiological functions and responses, and can be characterised using lipidomics, a specialised branch of metabolomics. As the end product of cellular activity, metabolites provide direct insight into physiological conditions, environmental response, and species interactions. At the ecosystem level, metabolomics can reveal how organisms contribute to food webs and biogeochemical cycles, including carbon (see Box 3.6.) and nutrient fluxes in the Ocean (Kujawinski, 2011).

Metabolic processes are highly sensitive to environmental factors such as temperature, light, and nutrient availability, therefore metabolomics can link organismal function to ecosystem dynamics by quantifying metabolites that mediate carbon and nutrient exchange among organisms and across food webs (Moran et al., 2016). Additionally, metabolomic "fingerprints" can be used to characterise species and communities, including cryptic taxa, and explore bioactive compounds with biotechnological potential. Metabolomics provides a direct and integrative snapshot of biochemical activity, capturing the combined effects of gene expression, enzyme function, and environmental conditions (Patti et al., 2012). Metabolomic profiles are closely linked to phenotype and physiological state, enabling detection of rapid responses to environmental change that may not be apparent at other molecular levels. Metabolomic analyses reveal active metabolic pathways, chemical signalling, and molecular exchanges (Moran et al., 2016). These strengths make metabolomics particularly powerful in marine systems, offering unique insights into e.g. microbe-molecule interactions and ecosystem-level biogeochemical function (see Box 3.6.).

[96] Metabolic processes are sets of chemical reactions carried out by living organisms that sustain life by producing energy and building cellular components. Metabolic processes within cells or organisms convert one molecule into another, forming ordered and connected metabolic pathways.

Box 3.6. Understanding microbial metabolites in the marine carbon cycle using metabolomics



Credit: Olivier Dugonay, Ifremer (CC BY 4.0).

Microbial metabolites in phytoplankton play an important role in marine carbon cycling.

Microbial metabolites connect biological activity with marine carbon cycling, acting as dynamic intermediates between photosynthesis, organic matter release, and remineralisation. Moran et al. (2022) showed that although microbial-derived metabolites constitute only a small fraction of the total dissolved organic carbon pool, they cycle rapidly and are central to the Ocean's role in global biogeochemistry. The study showed that one-quarter of photosynthesis-derived carbon can pass through a suite of short-lived metabolites generated by marine phytoplankton, heterotrophic bacteria, grazers, and viruses, linking primary production to microbial transformation processes.

Metabolomics can detect and quantify a diverse array of organic molecules in seawater, providing information on microbial activity and carbon transformation. These molecular fingerprints determine how microbial communities respond to environmental variability and how metabolite fluxes influence carbon turnover rates (Moran et al., 2022). Integrating metabolomic data with microbial community composition and environmental measurements advances functional understanding of the marine carbon cycle. It provides a mechanistic basis for linking microbial ecology with ecosystem-scale carbon flows, informing models of carbon sequestration and potential responses to climate change. Metabolomics can reveal complex biochemical pathways that traditional bulk measurements cannot resolve, providing information to improve marine biogeochemical models and their representation of Ocean carbon cycling, supporting assessments of the Ocean's role as a carbon sink, and strengthening the evidence base for ecosystem monitoring and climate-related marine policies.

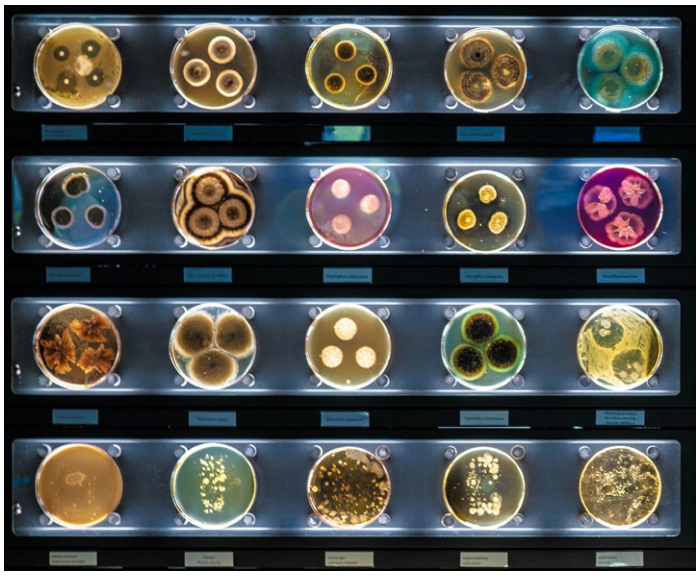
Metabolomic fingerprints of samples of individual organisms (tissue samples or cultures) or from environmental samples can be generated through proton Nuclear Magnetic Resonance[97] (NMR) spectroscopy or High-Performance Liquid Chromatography Mass Spectrometry[98] (HPLC-MS), or a combination of both. The results are compared against reference metabolite databases (e.g. BMRB[99], MassBank[100], METLIN[101], GNPS[102]) to identify the metabolites and estimate their abundance.

Metabolomics has not advanced at the same pace as genomics and transcriptomics, largely due to substantial analytical and interpretive challenges (Patti et al., 2012). Metabolites exhibit complex chemical structures and concentrations, and analyses rely on sophisticated and costly instruments, coupled with complex computational workflows (Table 3.1.). In addition, metabolite identification remains a major challenge due to incomplete reference databases and the difficulty of distinguishing between structural isomers[103] (Kujawinski, 2011). These limitations constrain throughput, reproducibility, and cross-study comparability, particularly in marine systems. Furthermore, metabolites detected in environmental samples cannot generally be directly attributed to the species or taxa that produced them. Establishing their biological origin therefore requires additional approaches, such as linking metabolomic data with genomic, transcriptomic or proteomic data (see Section 4.4.), or analysing metabolites from organisms grown under controlled conditions.

3.3.3. Reference databases

Reference databases are curated digital libraries of information about biomolecules that are used to determine the function and taxonomic origin of sequences by matching new sequences to known species and functions. Metabarcoding databases of specific gene regions (see Table 3.2.) used for taxonomic identification are incomplete, however those for vertebrates (e.g. fish) are more complete

97 Nuclear Magnetic Resonance is a technique to determine the structure and properties of molecules.
98 High-Performance Liquid Chromatography Mass Spectrometry is a technique to separate, identify and measure compounds in a complex mixture.
99 https://bmrb.io/
100 https://massbank.eu/MassBank/
101 https://metlin-nl.scripps.edu/landing_page.php?pgcontent=mainPage
102 https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp?redirect=auth
103 Isomers are molecules that have the same chemical formula but a different arrangement of atoms.



Credit: © Yann Caradec, Micropia Museum Amsterdam (CC BY 4.0).

Reference databases are very incomplete. For microorganisms they contain mostly those that are easy to grow in the laboratory, and many environmentally important organisms are underrepresented.

than those for invertebrates. A recent multinational effort to develop a curated barcode reference database for large organisms living at the seafloor in the North Sea determined that 2,514 species had been previously identified during morphological surveys (Christodoulou et al., 2025). However, barcodes were available for only 732 of these, with just 29% of all taxa covered, illustrating the scale of the work required to complete reference databases even in well studied regions. Building and maintaining such reference libraries is a substantial and long-term undertaking, requiring sustained investment in taxonomic expertise, specimen collection and identification, molecular characterisation, data curation, and interoperable database infrastructure. A federated system of taxonomically curated eDNA reference libraries (Rimet et al., 2026) is required, together with the technical architecture (i.e. the technical design, initial nodes, and software connectivity) and sustained investment to establish, connect, and populate these libraries.

Metagenomics, metatranscriptomics and metaproteomics analyses also depend on reference databases. Across all taxa, these databases are very incomplete. For microorganisms, these databases contain mostly fast-growing microbes that are easy to grow in the laboratory, and human pathogens, likely underrepresenting environmentally important taxa. As a result, the functions and taxonomic origins of many newly discovered genes remain unknown. For example, fewer than 10% of metatranscriptomic data can be assigned to a species, and many gene functions are unknown (Paoli et al., 2022).

In order to develop more reliable reference databases, it is necessary to generate DNA barcodes, genomes, transcriptomes and proteomes from a larger diversity of accurately identified species from taxonomic groups that are not yet well represented in these databases. Biomolecular approaches deployed on individual cells, tissues, or specimens to study their genome, gene expression, and epigenetic modifications provide in-depth genetic knowledge of species crucial for improving reference databases and gaining a better understanding of species relationships, functional roles, and evolutionary processes. In addition, reference databases can be improved by accelerating the experimental validation of gene functions, for example through advances in genetic approaches such as RNA interference[104] or CRISPR-Cas9 gene knockout[105]. These techniques allow researchers to selectively reduce or disrupt the activity of a specific gene and then observe the resulting biological effect, providing experimental evidence for its function. Such information can improve the annotation of encoded proteins and strengthen the link between taxonomy and function, thereby improving the interpretation of meta-omics data.

3.3.4. Bioinformatics and modelling for generating knowledge from data

The need for routine biomonitoring is driving innovations in bioinformatics and modelling, particularly through the use of environmental genomics and Machine Learning (ML) (Cordier et al., 2019). Bioinformatic tools are becoming more standardised, with new tools like eDNAmap[106] helping researchers compare biodiversity across different marine regions (Inoue et al., 2026). In addition, the scientific community is increasingly developing new numerical tools to integrate biomolecular data with other knowledge to improve understanding and prediction of marine ecosystem processes. These tools are stored and standardised in dedicated portals such as Europe's distributed infrastructure for life science data[107] (ELIXIR).

Multi-omics (see Section 4.4.) can be used to identify patterns across ecosystems and potential underlying mechanisms. For example, models of community-wide genomes can uncover patterns, such as amino acid and vitamin exchanges, that shape microbial communities in the upper sunlit part of the Ocean (Giordano et al., 2024). Data can also be converted into co-occurrence networks, which can help assess ecosystem structure, compare ecosystems, and predict ecosystem stability. For example, when applied to microbial ecosystems, plankton co-occurrence networks constructed from Tara Oceans metagenomes have demonstrated that specific community configurations

[104] RNA interference is a process by which small RNA molecules inhibit gene expression.
[105] CRISPR-Cas9 gene knockout is a genetic technique that uses the CRISP-Cas9 system to disrupt a specific gene, allowing researchers to determine its function by observing the effects of removing or disabling it.
[106] https://github.com/jun-inoue/eDNAmap
[107] https://elixir-europe.org/

and a small set of bacterial and viral genes strongly predict carbon export variability in oligotrophic regions (Guidi et al., 2016). In addition, global data on the co-occurrence and interaction of plankton species, integrated with modelling on the environmental conditions in which a species can live (i.e. niche modelling), can predict vulnerability of specific Ocean biomes to environmental change (Chaffron et al., 2021).

Computational methods are beneficial as they foster interdisciplinary research, using marine genomics to benefit a wide range of stakeholders. However, they face important limitations. They are mostly descriptive and based on correlation, meaning that their interpretation requires mechanistic understanding, making it difficult to identify general biological laws. In addition, correlative models cannot be used for prediction in new environmental conditions such as marine heatwaves (Muhling et al., 2020) and eDNA-based co-occurrence networks often do not directly reveal trophic interactions. In addition, these techniques rely on compositional data that represent relative parts of a whole (e.g. -omics data where a percentage of sequence reads is interpreted as a percentage of an organism within a community). Correlation analyses are very sensitive to these kinds of data and require strict standardisation to be comparable (see Section 5.2.).

As the costs of -omics data generation continues to decline, the bottlenecks are now primarily downstream data processing, storage, bioinformatic analyses and interpretation. These require substantial computational infrastructure, specialised expertise, and increasingly AI-assisted analytical workflows, making data analysis a growing component of the overall cost of translating -omics data into actionable knowledge (see Section 4.5.).

3.4. Recommendations

Recommendations for research funders:

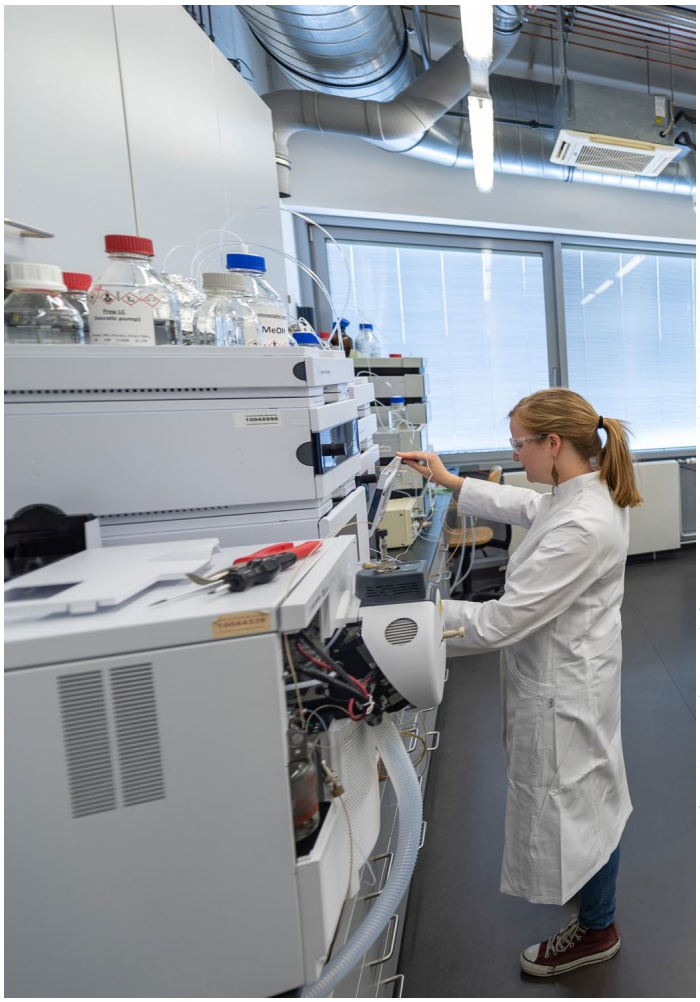
(3.1) Fund research to investigate DNA degradation kinetics, including temporal stability under different environmental conditions, the distances over which DNA is transported, and its sinking rate;

(3.2) Fund the further development of the architecture for a federated system of reference databases that are independently curated and maintained. Promote adherence to the principles of metadata publishing, quality control, and taxonomic curation, connected within a common network enabling information exchange, interoperability, and findability as envisaged by eDNAqua-Plan;

(3.3) Fund collection and sequencing campaigns to complete eukaryotic coverage for European waters for DNA barcode databases to facilitate metabarcoding and eDNA biodiversity analyses in support of MSFD, WFD, Nature Restoration Regulation and MPA monitoring;

(3.4) Fund research to generate genomes, transcriptomes and proteomes from a large diversity of accurately identified species from taxonomic groups that are not yet well represented in databases. This is required to develop more reliable and complete reference databases and increase the value and interpretability of existing and future biomolecular datasets; and

(3.5) Fund research to characterise genes of unknown function from existing and new model organisms to develop more reliable and complete reference databases of gene function.



Credit: © MARUM, Zentrum für Marine Umweltwissenschaften, Universität Bremen; V. Diekamp (CC BY 4.0)

Metabolomics laboratory.

4 Emerging technologies, and research and development needs

4.1. Summary

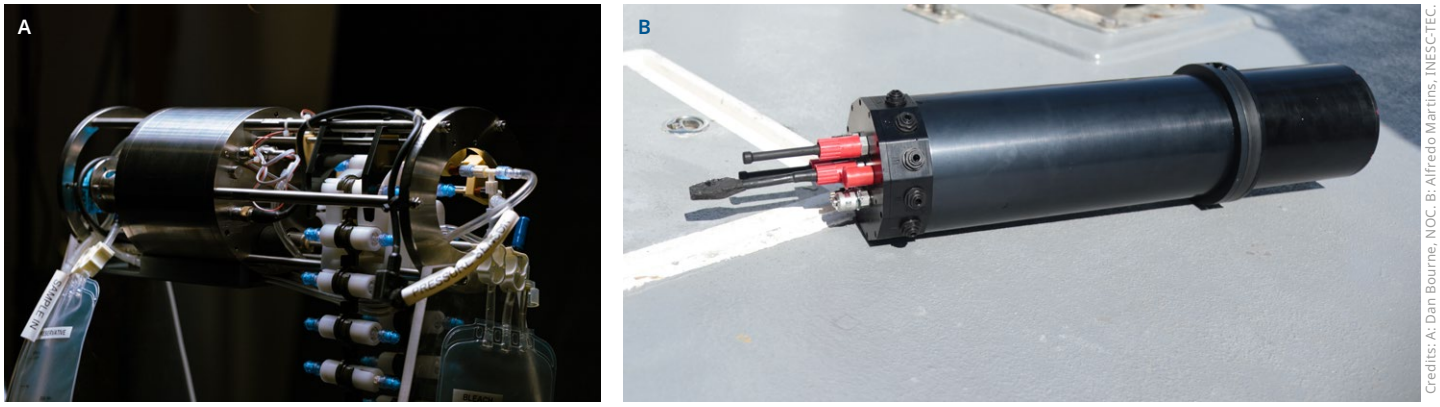
This chapter describes emerging developments in marine -omics that are delivering biodiversity observations at increasing temporal, spatial and taxonomic resolution, including high-throughput sampling and analytics from the field, to the lab, to the server (see Figure 4.2.). This chapter also discusses multi-omics, data analytics, and modelling, all of which will enhance marine biodiversity observing and interpretation, expediting the vision of a global -omics observatory and enhancing applicability for policymakers.

4.2. Deployable autonomous technologies

Deployable autonomous technologies include samplers that are used to collect water, sediment, or biological specimens, and sensors that detect and measure biomolecular signals. These are used in the field for automated, high temporal- and spatial- resolution measurements of marine life.

4.2.1. Autosamplers and passive samplers

The workflow for environmental DNA (eDNA) and -omics sampling and sensing is time-consuming since filtering water to concentrate eDNA can take hours. Deployable autosamplers have therefore been developed to save time by pumping and filtering water *in situ* at predefined times, depths or locations, enabling high-throughput sampling of eDNA, RNA and proteins across both fixed and mobile platforms (Figure 4.2.), from the surface to the deep sea. An array of autosamplers is now in development, and some are commercially available, and they address various use cases (Yamahara et al., 2025). Among these, the Robotic Cartridge Sampling Instrument[108] (RoCSI; NOC, United Kingdom; Figure 4.1A), the *in situ* autonomous biosampler (IS-ABS; INESC TEC, Portugal; Figure 4.1B; Ribeiro et al., (2019)), inDepth[109] (Applied Genomics, United Kingdom) and MicroSense[110] (Observatoire Océanologique de Banyuls-sur-Mer, France) were invented in Europe.



Credits: A: Dan Bourne, NOC. B: Alfredo Martins, INESC-TEC.

Figure 4.1. Two examples of European autosampler technologies for *in situ* eDNA concentration and preservation.
A: The RoCSI (left) from the National Oceanography Centre, is capable of capturing 45+ samples in its standard configuration, and is operational to 6,000 m. **B:** The IS-ABS from INESC-TEC is smaller, and can collect 16 samples to 150 m.

[108] https://mclanelabs.com/rocsi-landing/
[109] https://appliedgenomics.co.uk/sample
[110] https://www.embrc.eu/from-lab-to-market-transforming-marine-monitoring-with-biosensors/

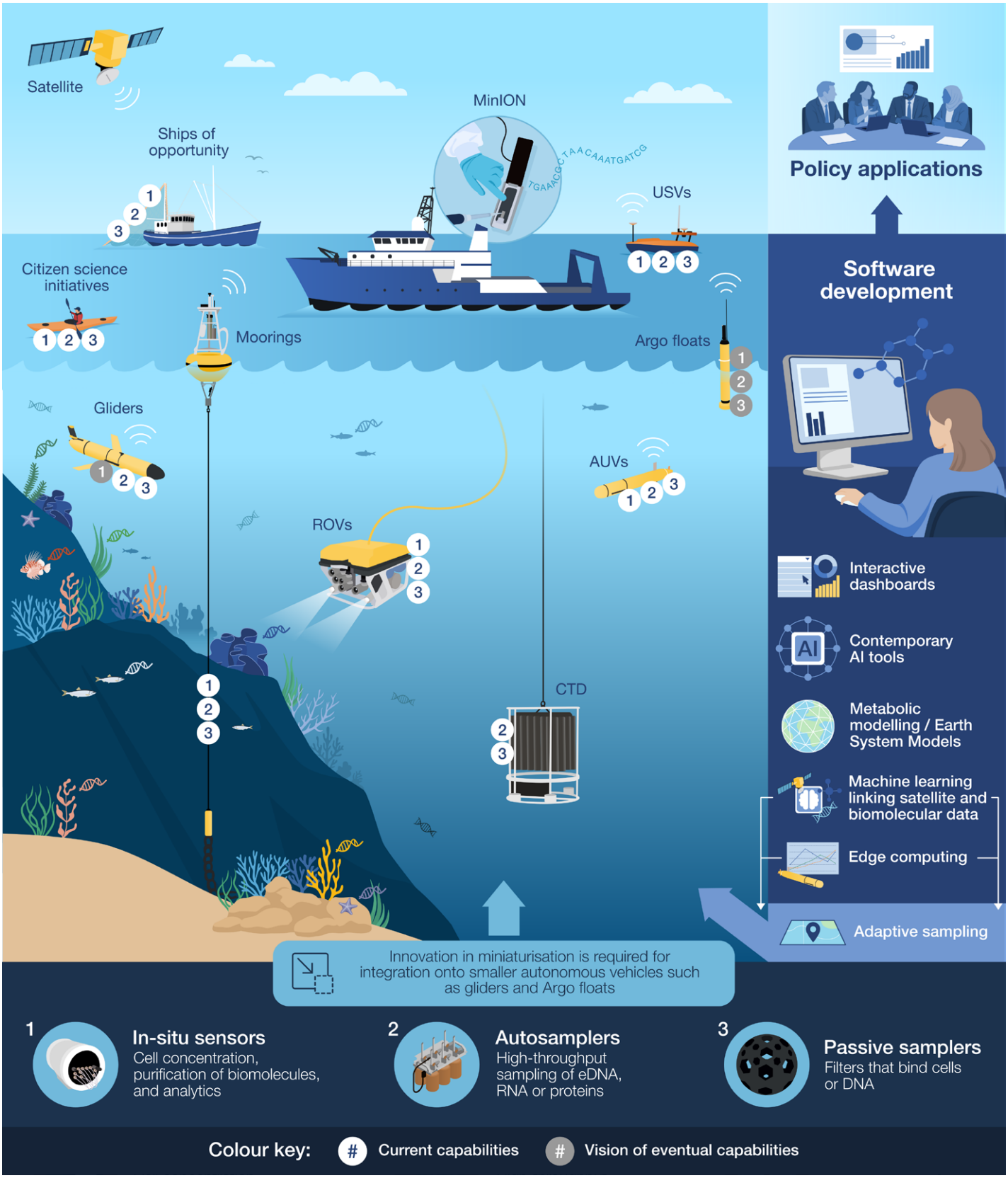


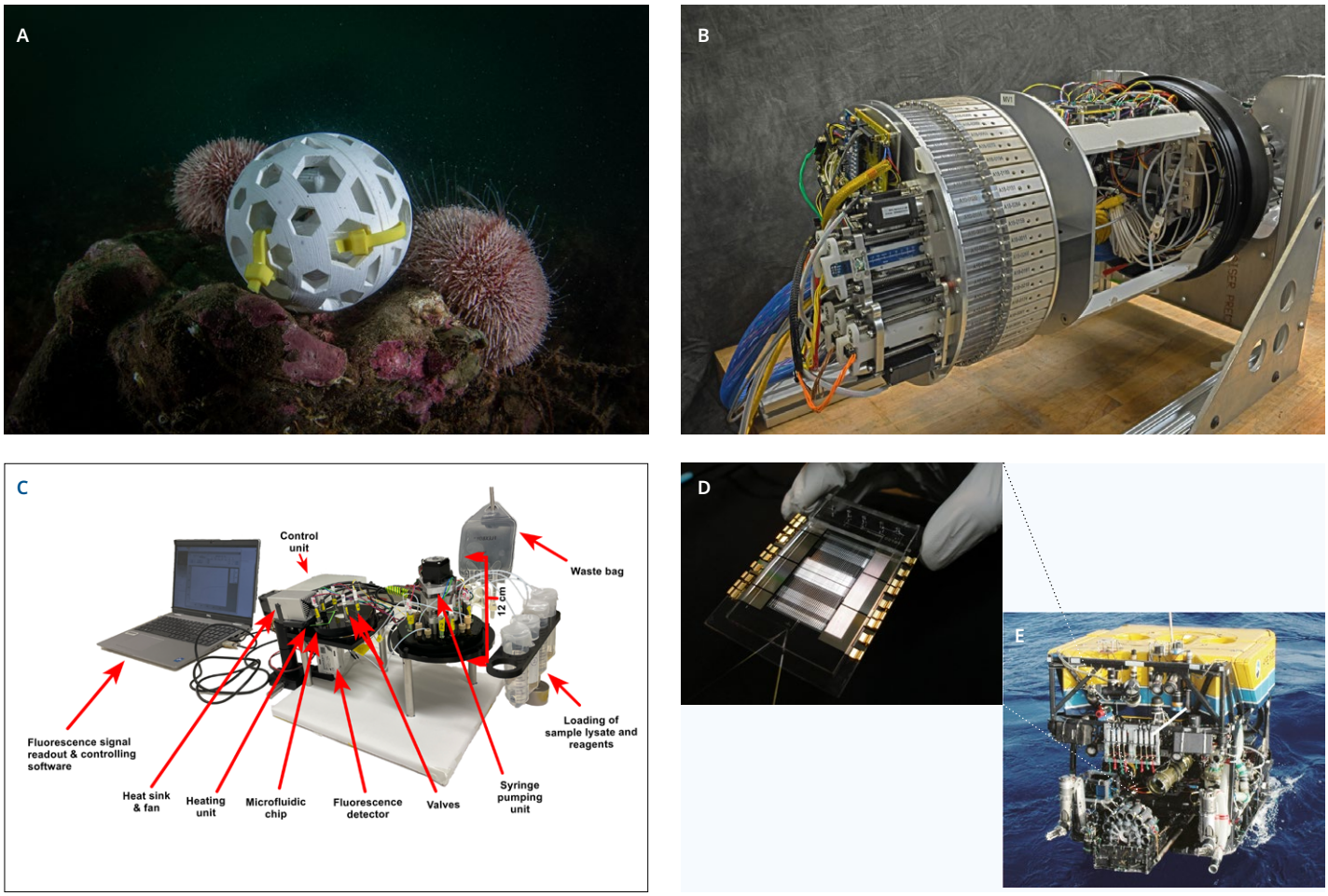
Figure 4.2. Overview of the vision for an automated -omics to knowledge pipeline. *In situ* sensors (1), autosamplers (2) and passive samplers (3) for biomolecular observing are identified for each platform. White numbers correspond with technologies that have already been demonstrated, and grey numbers are visions for / upcoming sensor-platform development and integration. The vision of data flows is shown on the right, where recently-captured data can inform vehicle navigation or operations i.e. "adaptive sampling". A parallel, important stream of development ensures collected data are repository- and AI-ready, which would enable harvesting into public-facing dashboards and models at appropriate timescales for policymakers.

Advances in passive eDNA samplers are also being made that accumulate biomolecules through adsorption using a filter directly placed in seawater. These take advantage of the passive movement of water over the filters without using pumps that are required for autosamplers. There are many variations of passive samplers that simplify eDNA acquisition for use on ships of opportunity[111] or in remote environments, where conventional instrumentation is impractical. For example, the Metaprobe (Figure 4.3A) is an agile, low-cost passive eDNA sampler that has been used extensively to identify marine vertebrates and other biodiversity (Maiello et al., 2022) and can be attached to fishing nets and other deployable devices.

Both autosampling and passive sampling can be conducted using Conductivity, Temperature and Depth (CTD) rosettes, Autonomous Underwater Vehicles (AUVs), Remotely Operated Vehicles (ROVs) and moorings. Many of these samplers are commercially available and require minimal training effort, while passive samplers can additionally be 3D printed by users. The range of autosamplers and passive samplers is expanding. Such simplified sampling approaches, combined with citizen science initiatives, enable large-scale biodiversity characterisation. Data generated from these efforts are increasingly shared through interactive platforms, such as the Integrated Coastal Observatory in British Columbia, Canada[112].

4.2.2. *In situ* sensors

In situ sensing allows biomolecular measurements to be taken in the Ocean without the need to bring samples back to the lab. Autonomous *in situ* sensors process samples to detect or quantify specific genes or transcripts (DNA or RNA). This requires sample filtration to concentrate cells, breaking open the cells, and molecular analytics, which is typically optical and based on fluorescent dyes or probes (see Chapter 3). Some of these sensors can also save a subset of samples to return to the lab.



Credits: A: Bob Anderson; B: © Todd Walsh, MBARI; C Alrefaey et al., (2026) (CC BY 3.0); D: Tatsuhiro Fukuba © JAMSTEC; E: © JAMSTEC cruise NT08-11, Dive: #853, ROV: Hyper-Dolphin, Photographer: Cruise participant (Tatsuhiro Fukuba)

Figure 4.3. Scaling up of *in situ* molecular technologies is occurring across parallel development pathways, ranging from broad-scale application of low-cost devices to development of small *in situ* sensing instrumentation. **A:** The Metaprobe is an affordable, open source, 3D printable cage to mount filter membranes internally for eDNA capture; **B:** The Generation 3 Environmental Sample Processor (ESP3) high-resolution sampling and extraction device is an *in situ* sampler and is undergoing development for *in situ* DNA and RNA analytics; **C:** An adaptation of the 6,000 m depth rated lab-on-a-chip wet chemical sensor for genetic sensing. The components shown have been demonstrated in the laboratory, and are in development for *in situ* sensing; **D & E:** The Integrated *in situ* genetic analyser (IISA-Gene) lab-on-a-chip PCR device has quantified rRNA and functional genes from an ROV in the deep sea.

[111] Ships of opportunity are vessels such as commercial ships, ferries or fishing boats, that can be used to collect environmental samples or observations during their normal operations.

[112] https://ico.hakai.org/map?group=All&site=null&species=null&startDate=2023-12-01&endDate=2025-12-01

Many autosampler developers are simultaneously developing integrated sensors. For example, the Environmental Sample Processor[113] (ESP), developed by the Monterey Bay Aquarium Research Institute[114] (MBARI), is a versatile and efficient platform capable of multiple autonomous molecular analytics in near real-time (Scholin et al., 2017). The Generation 2 ESP is capable of eDNA and microbial filtration and preservation at a user-determined frequency, for up to 144 samples over weeks- or months- long deployments. It's deployed on dedicated moorings routinely by the National Oceanic and Atmospheric Administration (NOAA) in the USA as an early-warning system for harmful algae, for monitoring threatened and invasive species, and faecal indicator bacteria (e.g. Yamahara et al., 2015). To miniaturise and thereby increase agility for deployment, the Generation 3 ESP (see Figure 4.3B) performs sample acquisition and preservation, and is in development for microfluidic DNA analysis capabilities. This sensor will be small enough to be deployed for biodiversity surveys on long-range vehicles such as AUVs, gliders and Unmanned Surface Vehicles (USVs) (Yamahara et al., 2019). It has been demonstrated for *in situ* harmful algae toxin detection in real-time on a USV (Ussler et al., 2024).

Miniaturisation is key to broadening the range of platforms that *in situ* sensors can be deployed on, thereby increasing their potential range of applications in research and monitoring. Lab-on-a-chip[115] systems (see Figure 4.3C & 4.3D) have also been developed as miniaturised *in situ* deployable molecular sensors and demonstrated by deployment on a Remotely Operated Vehicle (ROV; Figure 4.3D-E) exploring deep-sea hydrothermal vents, employing techniques such as quantitative PCR (qPCR) (Fukuba & Fujii, 2006). Similarly, DNA extraction and analytical capabilities have been developed for other lab-on-a-chip platforms and these are undergoing further development for cell concentration (Alrefaey et al., 2026).

Across all sensor types, there is a need to develop easier-to-use hardware and software to broaden uptake. Edge computing (i.e. localised data processing on *in situ* sensors) is cost-effective, allows quicker response times and real-time decision-making. When provided with real-time contextual information about the specific target molecule and organism, molecular sensors can potentially estimate abundances and therefore trends for many marine organisms *in situ*.

4.2.3. Adaptive sampling on autonomous platforms

Adaptive sampling focuses biomolecular sampling to specific features of interest. This involves uncrewed vehicles or platforms that autonomously decide where and when to collect samples, using algorithms fed with predetermined oceanographic parameters such as bathymetry or chlorophyll concentration (Zhang et al., 2019). This is important because samplers and sensors have a restricted number of samples that they can acquire. Therefore, decisions on where and when to sample based on existing knowledge are required for long-endurance (e.g. weeks-long) missions. Additionally, both sensors and samplers require cell concentration using (e.g. submersible pumps), which requires a lot of power, which is currently the limiting factor in the length of long-duration deployments. One significant exception to this are eDNA technologies deployed on USVs capable of solar power generation and wave propulsion.

To broaden the capabilities of autonomous biodiversity surveys, algorithms can be further developed using data from *in situ* sensors or the vast array of satellite-derived Ocean surface data, available in near real-time. Algorithms can be used to ensure that satellite and sensor-based observations are integrated into plans for sampling campaigns. This enables -omics data to be collected that are relevant to short-lived biological features or oceanographic features such as fronts and eddies that are important for marine biogeochemistry but challenging to capture in large-scale oceanographic campaigns.

4.2.4. Technological development for deployable technologies: A future outlook

In situ technologies have the potential to transform marine biodiversity observing across the three-dimensions of the Ocean, in the same way as satellite observations have improved our ability to monitor the surface Ocean. The range of deployable samplers available enables a diversity of scientific applications, with no single sampler fitting all use cases. Many are commercially available, enabling broad implementation, but further innovation is required.

Miniaturisation of all deployable technologies, including samplers and sensors, will ensure integration onto smaller autonomous vehicles such as gliders and Argo floats. Innovation to ensure the expansion of high-throughput, high-sample-volume autosamplers and high-throughput analytical capabilities is required. Smaller, easy-to-use sensors that measure DNA using electrical signals (i.e. electrochemical sensors) exist for targeted species, and can be diversified for shorter-term campaigns or developed further to increase robustness (Metfies et al., 2005). Innovation is required for more diverse passive sampling capabilities (i.e. to eliminate the need for power-hungry pumps in electronic multi-samplers for long-term deployments), and analytical tools to measure multiple target genes or transcripts simultaneously such as multiple-probe-based (Shilova et al., 2014) or miniaturised, multiplexed (i.e. measuring several target molecules simultaneously) single-use microfluidic[116] chips. The development of microfluidics to prepare samples for nanopore sequencing[117] is also an area primed for future development. Work is ongoing at Monterey Bay Aquarium

[113] https://www.mbari.org/technology/environmental-sample-processor-esp/
[114] https://www.mbari.org/
[115] Lab-on-a-chip are miniaturised devices that integrate lab functions onto a small chip.
[116] Microfluidics is the science and technology of manipulating very small volumes of fluids through microminiaturised channels.
[117] Nanopore sequencing detects changes in current as individual nucleic acid molecules pass through a nano-scale pore.

Research Institute and Johns Hopkins University Applied Physics Laboratory[118] in the USA to purify and prepare genomic DNA from marine environmental samples for microfluidic sample preparation and long-read sequencing using the Oxford Nanopore (Yamahara et al., 2025).

To accelerate work and minimise replication, autonomous sensor and sampler instrumentation developers coordinate internationally and meet regularly through initiatives such as the Marine Omics Technology and Instrumentation (MOTI) community[119]. Newly developed fit-for-purpose samplers need to be validated to check that they deliver comparable datasets. At the moment, this is done by in-person intercomparison exercises, such as those by the MOTI community to evaluate intercomparability of DNA datasets. There is a need to develop commercially available standards, such as the Zymobiomics[120] microbial community standards, but for a diversity of taxa, to enable new technologies and capabilities to be validated for distinct taxa without the need for in-person intercomparison exercises. These would additionally help the global eDNA and -omics communities to standardise and quantify specific taxa and their trends using molecular data, which is needed for marine management (see Section 5.2.).

These transformative technologies are only as strong as the data pipelines and repositories that support them, and technology developers need to standardise and align the way that data and metadata are described, structured, and formatted as part of the technology development work (see Section 5.3.). Expanding the use of deployable technologies combined with low-cost and portable sampling and sequencing capabilities will accelerate the generation of sequence data at unprecedented scales. However, this will require fit-for-purpose data flows and mature visualisation tools to ensure maximum impact and enable the resolution of biodiversity trends.

4.3. Emerging molecular laboratory technologies: Long-read and miniaturised sequencing

High-throughput sequencing delivers increasing accuracy at decreasing cost. Sequencing methodologies have undergone three major transformative stages (i.e. first generation Sanger sequencing, followed by high-throughput sequencing and single-molecule nanopore sequencing; see Section 3.3.), with each delivering an exponential acceleration in capability. Nanopore sequencing is a relatively new technology for long-read sequencing of DNA and RNA, released in the early 2010s by Oxford Nanopore Technologies. It is typically used for genome sequencing, but full-length sequencing of specific genetic markers is also possible. The portability of the technology can give researchers insight into biodiversity and community composition while in the field, allowing real-time strategic expedition planning and adaptive management.



Credit: Oxford Nanopore Technologies.

The standard MinION flow cell is a key component of Oxford Nanopore Sequencing. Its portability has enabled its application in remote and under-resourced regions including use on research vessels in the open Ocean, polar environments, and on remote islands.

For example, it can target specific communities for experimentation in real-time, flag an area potentially inhabited by threatened or endangered species, and improve timely management of waterborne threats. Longer eDNA reads from specific taxonomic groups of interest would deliver population level insights in a non-destructive manner.

Nanopore sequencing uses a radically different method of sequencing based on an engineered protein pore, allowing miniaturised, ultralong-reads that increase accuracy and allow high-throughput sequencing at low cost. The standard MinION flow cell is a small, handheld cartridge with thousands of nanopores, and is a key component of Oxford Nanopore Sequencing. This has enabled scaling the technology and its application in remote and under-resourced regions including on research vessels in the open Ocean, polar environments, and on remote islands. Nanopore sequencing at sea enables the researcher to leave the ship with data, rather than with hundreds of samples that take weeks to months to process. However, standardised analytical pipelines are required to scale-up its use at sea (see Section 5.2.). The low cost has also accelerated its use in less developed countries, and it is widely used in capacity development efforts.

In general, standard practices evolve over time as protocols become obsolete when better technologies are released, and older laboratory supplies are often discontinued by the few companies dominating the market. Therefore, significant periods of overlap and intercomparison are required when adopting new sequencing methods to ensure users are clear on how to interpret data and that laboratories using previous methodologies continue to collect valuable data.

[118] https://www.jhuapl.edu/

[119] https://sites.google.com/mbari.org/moti-workshop/home

[120] https://zymoresearch.eu/collections/zymobiomics-microbial-community-standards?srsltid=AfmBOopaQKdI0sdcCg9xu2mHyBkHE5wrfMLv4wGdFJCHmZLI2DfFR6o6

4.4. Multi-omics and scaling augmented observatories

As described in Chapter 3, the choice to use metagenomics, metatranscriptomics or metaproteomics depends on the scientific and policy question, and each has complementary advantages. However, multi-omics is also possible, which integrates data from two or more -omics technologies to provide an accurate, comprehensive and holistic perspective of ecosystem processes. As costs of -omics analyses decrease (see Chapter 3), multi-omics studies will proliferate to deliver holistic understanding of organisms' responses to changing environments and the change in function that results.

For example, the simultaneous collection of metagenomic and metaproteomic data provides a sample-specific protein database to help identify more molecules from metaproteomic data by matching them to their coding genes. Metagenomics allows understanding of species presence and enables ecological insights, while also enriching interpretations and forecasts of protein production and function changes under different environmental conditions. The advantages of multi-omics have been demonstrated in corals where the combined use of transcriptomics, and proteomics and metabolomics gave a clearer understanding of the physiological impact of environmental stressors (Williams et al., 2023).

At the vanguard of next-generation of Ocean monitoring at scale is the Traversing European Coastlines[121] (TREC) programme. Leveraging -omics best practice, the campaign sampled 115 land–sea transects along European coastlines between April 2023 and July 2024. By integrating meta- and single-cell -omics, DNA barcoding, advanced microscopy and detailed environmental measurements, TREC emphasised the need for augmented observatories across European coastlines for characterising biodiversity from viruses and bacteria to algae and animals. TREC is a complete observatory model that should be adopted more broadly. The resulting molecular, morphological and environmental data will provide an unprecedented pan-European census of coastal ecosystems and their responses to natural and human-induced environmental gradients.

4.5. Computational methodologies to enrich data interpretation

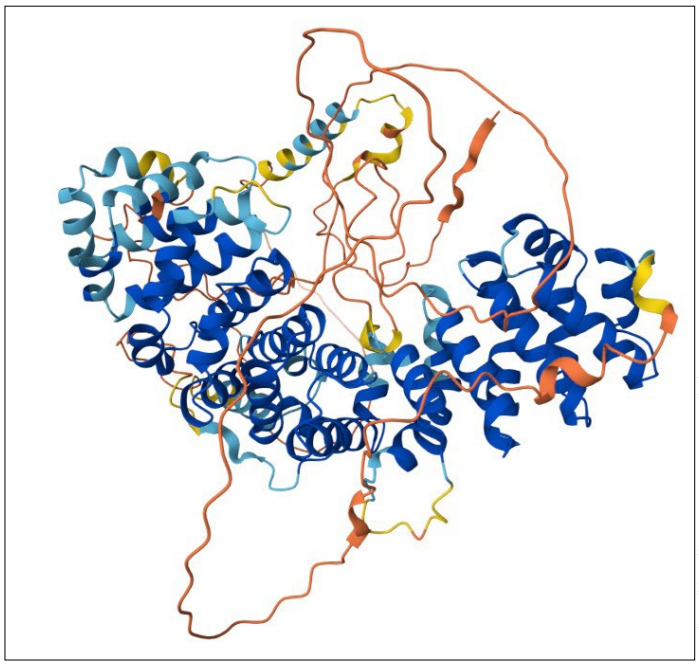
Emerging technologies are accompanied by new numerical methods. In particular, Machine Learning (ML) has already been used to help analyse -omics data and is anticipated to be further deployed. For example, ML can be used to predict trends by associating -omics data with other types of knowledge. For example, the integration of ML with satellite imagery is becoming a standard method for using -omics data to predict biological functions and biodiversity across the global Ocean. Predictions have been made of global plankton community distribution from satellite data with approximately 67% accuracy (Kaneko et al., 2023). In addition, this approach has been used to improve habitat predictions (Mugo & Saitoh, 2020) and produce predictive maps of Antarctic plankton biological interactions (Grillo et al., 2024). In the near future, further generalisation of these approaches is expected. ML will be used to align high-resolution environmental data (e.g. physics, optics) with biomolecular data to identify global-scale trends, such as bioprovinces. Identification of such relationships has been used to project trends when biological data aren't available, providing potential guidance for future oceanographic campaigns. In particular, ML that links global sequence datasets with Earth observations from satellites has already resolved phytoplankton community structure and improved global estimates of primary production (El Hourany et al., 2024). A recent study that interpreted larger-scale Ocean observations is the standardised and highly automated modelling framework "Comprehensive Ensemble Pipeline for Habitat Modelling Across Large-scale Ocean Pelagic Observation Datasets" (CEPHALOPOD) (Schickele et al., 2025), which used ML to merge -omics and imaging data to make more accurate predictions of biogeography across multiple species, which was only possible using high-resolution, machine-readable datasets.

Marine -omics research has also benefitted from recent advances in Artificial Intelligence (AI) and the availability of large-scale datasets (see EMB Future Science Brief N°. 6 on Big Data in Marine Science; Guidi et al., 2020). Contemporary AI tools trained on extensive prior knowledge enable the extraction of novel biological properties from marine sequence data at high accuracy and scale. These tools can unlock the functional and biochemical complexity of marine ecosystems at large scales and high resolution.

For example, AlphaFold[122] is an AI programme that predicts protein structures within minutes with high accuracy. These data are available in the AlphaFold Protein Structure Database[123], which has massively expanded structural coverage of protein-sequences (Varadi et al., 2022). Marine metagenomic datasets have been explicitly leveraged to predict the three-dimensional structure of proteins and enable structural and functional predictions for previously uncharacterised protein families from Ocean microbes.

Other examples of AI tools automating biological discovery from -omics data are AlphaGenome[124]: a deep learning model that processes DNA sequences and advances their functional annotation (Avsec et al., 2026); DeepVirFinder[125]: used for the discovery of new viruses uses convolutional neural networks to identify lesser-known viruses in metagenomic data without reliance on reference databases (Mo et al., 2026); and CFM-ID

[121] https://www.embl.org/about/info/trec/expedition/
[122] https://deepmind.google/science/alphafold/
[123] https://alphafold.ebi.ac.uk/
[124] https://deepmind.google.com/science/alphagenome/
[125] https://github.com/jessieren/DeepVirFinder



Credit: Wikimedia Commons, BQUB24-Ebahr (CC BY-SA 4.0).

AlphaFold is a highly accurate AI tool that predicts protein structure within minutes.

4.0 framework[126]: used in metabolomics to improve compound identification (Wang et al., 2022). However, these tools are not specific to marine systems. Future improvements for marine systems will rely on access to marine -omics data (see Section 5.3.), the use of explainable AI[127] to guide future experiments, and new biological discoveries.

Among explainable AI approaches, metabolic modelling offers an explicable, mechanistically grounded way to study marine ecosystems. Recent advances in constraint-based programming[128] allow the reconstruction of genome-scale metabolic models (GEMs), where genomic data are converted into networks of biochemical reactions that define the metabolic capacity of an organism, a community, or an entire ecosystem (Sher et al., 2024). These GEMs rely on mechanistic simulations of thousands of reactions to predict metabolic traits and ecological niches, which allows the reconstruction of trophic networks, unveiling metabolic cross-talk among marine organisms, characterising the functional roles of microbiomes in association with their hosts, and linking molecular processes to biogeochemical cycles. The integration of GEMs into Earth system models such as NEMO-PISCES[129] was biologically validated and enables the explicit inclusion of genomically described organisms within predictive global frameworks (Régimbeau et al., 2025). This integration supports evaluation of the concept of metabolic engines[130] (Falkowski et al., 2008) and calls for its generalisation for ecosystem-wide molecular diversity to link cellular-scale metabolism to planetary-scale fluxes. The overarching goal of the BioGeoSCAPES[131] program is to validate such an approach. By embedding genomic and metabolic detail into Earth system simulations, metabolic modelling will enhance the accuracy and predictive fidelity of models, designing the next generation of -omics based Digital Twins of the Ocean (DTOs) advancing our capacity to understand and manage the molecular foundations of marine biogeochemistry (Tagliabue, 2023) and the potential role of evolution.

Finally, to understand and add value to the interpretation of large and complex -omics data, analytics and visualisation can be combined in dedicated platforms, such as Anvi'o[132]. The communication of biomolecular results through dedicated dashboards or Python notebooks can also ensure that policymakers and managers can replicate science-based insights (see Box 4.1.).

4.6. A vision for molecular biological applications for the Ocean

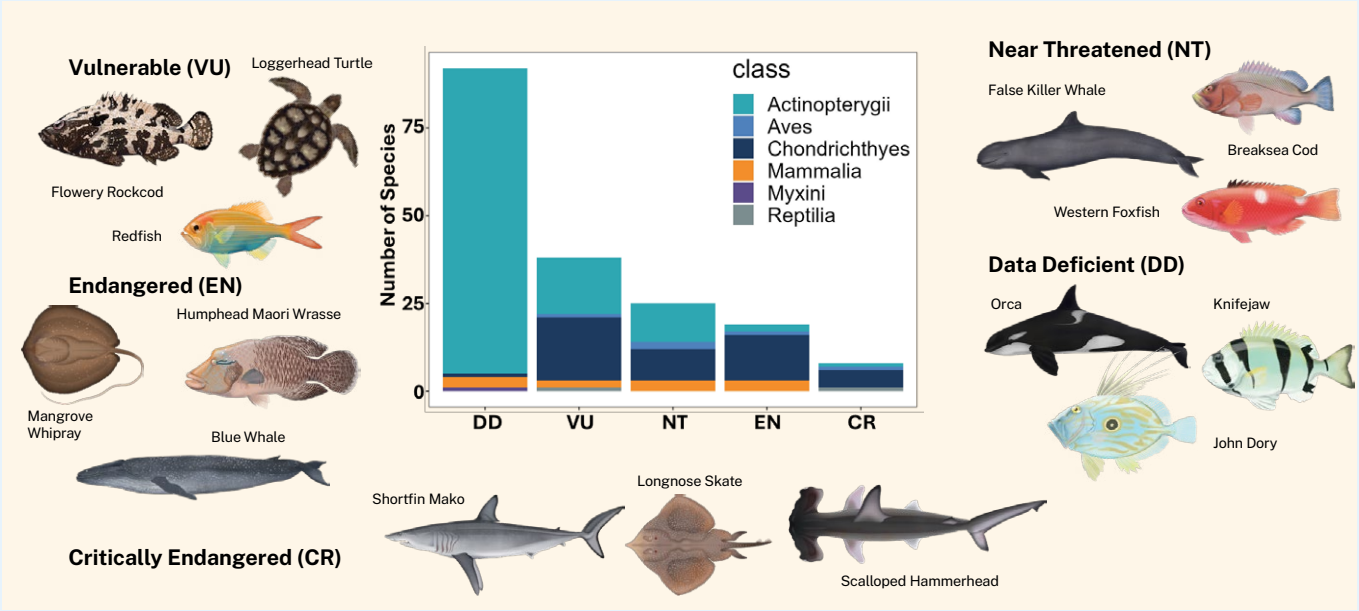
Individually, each aspect mentioned in this chapter (deployable technologies, laboratory techniques, multi-omics, data analytics, and modelling) will enhance marine biodiversity observations and interpretation, expediting the vision of a global -omics observatory, and enhancing applicability for policymakers. This will revolutionise our capacity to monitor, predict and respond to changes in marine biodiversity. Reducing costs and improvements in sequencing technologies would further expand the use of biomolecular technologies globally. Ultimately, expediting the flow of *in situ* data to the people who respond to it will ensure adaptive management, while accelerating data-led ecosystem and economic protection for the sustainable use of the Ocean.

A roadmap towards a vision for scaling-up biomolecular tools for Ocean research and management begins with the expansion of sensor development and analytics, especially adopting AI and *in situ* technologies, individually and together. Scaling low-cost technologies, such as passive sampling and citizen science, and the expansion of augmented observatories will result in large amounts of data that will overwhelm current repositories, which need to adapt and expand. Importantly, biomolecular sampling and analysis should be accompanied by data reporting by default. Dedicated funding to link technology developers and citizen scientists to data experts in order to supply AI-ready data with controlled semantics, vocabularies and reporting to shared repositories and dashboards, would ensure that the impact of these developments is fully realised (see Chapter 5).

[126] https://cfmid.wishartlab.com/
[127] Explainable AI is a set of tools and methods that allows humans to understand and trust the outputs created by machine learning.
[128] Constraint-based programming applies a set of rules to analyse and predict the behaviour of complex systems.
[129] https://nemo-ocean.eu/
[130] The concept of metabolic engines considers each organism as biochemical machinery whose interaction contributes to biogeochemical cycles.
[131] https://biogeoscapes.org/
[132] https://anvio.org/

The legacy of emerging biomolecular technologies and European sovereignty in Ocean observing can be achieved by incentivising stronger academia-industry collaboration for the development and commercialisation of eDNA technologies, and using European numerical Research Infrastructure (i.e. MGnify[133]) both as scaling platforms for European commercial technologies and as test beds for prototypes and iterative development, as recommended by the European Ocean observing community[134].

Box 4.1. The Minderoo Marine Parks Dashboard



The Minderoo Marine Parks Dashboard enables users to interrogate data regionally and to ask questions about the data, including species and their IUCN conservation status.

The Minderoo Marine Parks Dashboard[135] was developed to interrogate thousands of biomolecular samples from hundreds of sites across the 4.2 million km^2 of Australia and Commonwealth Marine Parks for species of interest. Australia's rich marine biodiversity is represented across the three billion sequence reads covering 2,217 marine vertebrate species, including 2,064 bony fishes, 94 sharks and rays, 31 mammals, 23 birds and three turtles. The dashboard can generate tables of species present in regions of interest (Raes et al., 2026) and AI is used to answer user questions about the data. The dashboard maps species based on IUCN conservation status, and over time, to understand trends based on their eDNA signatures across expeditions (see figure). The dashboard has a well-curated database associated with it, and further work from the Minderoo Foundation seeks to improve the accuracy of the underlying reference databases by sequencing mitochondrial genomes from all known families of marine vertebrates using taxonomically characterised museum specimens. However, there remain unidentified taxa from sequences that are present but unknown and the dashboard shows which families these sequences are related to. The dashboard was funded through a collaboration between the Minderoo OceanOmics programme[136] and Australian Marine Parks (i.e. both government and philanthropy).

[133] https://www.ebi.ac.uk/metagenomics/
[134] https://marcobolo-project.eu/wp-content/uploads/2026/05/Blue-Tech-Europe-Policy-Recommendations-2026.pdf
[135] https://marine-parks.minderoo.org/
[136] https://www.minderoo.org/resources/oceanomics/

4.7. Recommendations

Recommendations for research funders:

(4.1) Fund innovation in the development of low-cost, miniaturised deployable samplers and sensor technologies in order to accelerate their use at scale;

(4.2) Fund the expansion of the use of AI for faster sequence identification using well-curated databases, and to associate -omics data with other environmental parameters to predict planetary-scale trends;

(4.3) Fund research to incorporate plankton metabolism and molecular traits into Digital Twins of the Ocean and IPCC models to support spatial and climate scenario projections of -omics functional knowledge;

(4.4) Accelerate a European eDNA technology pipeline in the first instance, with other technologies to follow, by strengthening support for technological development and high-impact demonstrations, while incentivising commercialisation;

(4.5) Provide dedicated funding to link technology developers, citizen scientists and data experts to standardise and align the description, structure and format of data and metadata to supply AI-ready data with controlled vocabularies and ensure reporting to shared repositories and dashboards that include advanced data visualisation tools; and

(4.6) Fund multi-omics research across diverse species to better link organisms to their functions, improve accuracy of models and spatial and temporal predictions, and explore new "-omics"-informed theories.

Recommendations for scientists:

(4.7) Perform intercomparison exercises with environment agency standards, to expand the adoption in the field of user-friendly portable technologies starting with underway autonomous and manually deployable eDNA sampling systems, saving costs and time at sea.

Recommendations for policymakers:

(4.8) Integrate information on biodiversity with other data sources (e.g. oceanographic data, remote sensing, taxonomy databases) in monitoring surveys for holistic marine ecosystem assessments.

5 Operationalisation, standardisation and FAIR data

5.1. Summary

This chapter outlines how operationalisation, standardisation and the adoption of FAIR (Findable, Accessible, Interoperable, Reusable) and open data principles enable the generation of knowledge products that can address marine policy needs at scale, and its integration in Digital Twins of the Ocean (DTOs). While focusing on the openness of the data, this chapter also frames the need for access and benefit sharing of genetic resources.

5.2. Operationalisation

Operationalisation transforms abstract biological concepts into measurable proxies. This is important because living systems are inherently complex and describing them requires tracking a large number of difficult to measure variables. For example, conceptually biological diversity is the entirety of all varieties of life, but that is hard to quantify in practice. While counting species is relatively straightforward, genetic and functional diversity, as well as the diversity of species interactions, cannot be directly translated into numbers. Simplified mathematical concepts, such as interaction networks, are therefore used to describe complex biological systems. Similarly, physiological processes and activity are difficult to assess as they involve complex biochemical processes. Researchers therefore measure properties as proxies, such as the expression of genes that encode enzymes responsible for specific reactions (Conesa et al., 2016). Gene expression itself is operationalised as the number of messenger RNA (mRNA) transcripts at any given moment in a cell (Wang et al., 2009). Such operationalisations are validated through peer-review and debate and are revisited when new technologies enable better observations of biological systems.

The operationalisation of biological processes is not perfect and may introduce errors and biases, requiring standardisation so they can be more consistent and comparable among researchers. For example, plankton covers multiple orders of magnitude in size, and abundance cannot be measured with a single lens or pore size, requiring standardisation of size fractions (Lombard et al., 2019). However, standardisation should match biological patterns and processes and not technological limitations. In practice, it is usually an iterative mixture of the two, where different experts discuss optimal standardisation. Initial procedures and definitions are shared, criticised, tested, and refined until they become functional, agreed, and verifiable. Operational costs also play a role. If the best observation technique is prohibitively expensive, the standard might be lowered to facilitate large-scale adoption. In contrast to operationalisation, standardisation is less dependent on scientific advances, and is usually enforced by an authoritative body, e.g. a standard consortium including the scientific community, policymakers, civil society and the economic sector.

5.3. Standardisation

The lack of formal standards currently limits the adoption of biomolecular methods into regulatory frameworks such as the MSFD and the WFD (Blancher et al., 2022). Standardisation is needed to make biomolecular data and methods more consistent and comparable to produce actionable results at scale. Standardisation encompasses all steps of biomolecular studies from sample collection, data analysis, and result interpretation. For example, standards may mandate the inclusion of a minimal level of sampling replication and negative (no-DNA blanks) and positive (DNA reference material) control samples in environmental DNA (eDNA) metabarcoding.

Standardisation also fosters international collaboration, as demonstrated by the nearly universal adoption of the FASTA (text-based) format[137] for storing DNA sequence data, which greatly facilitates global analyses.

Biomolecular observations are inherently complex because species and physiological processes are derived from molecular signatures. These molecules need to be extracted from cells and tissues, and may degrade during preservation, handling or storage. Each step of a biomolecular method is prone to biases (see Chapter 3), and it is critical to define

[137] https://www.ncbi.nlm.nih.gov/genbank/fastaformat/



Credit: © Hanneloor Heynderickx.

Standardisation makes biomolecular data and methods more consistent and comparable to produce actionable results at scale.

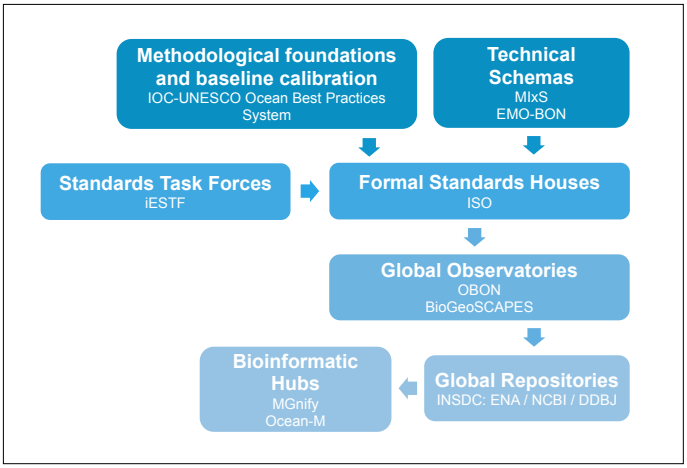


Figure 5.1. The standardisation landscape in marine biomonitoring. For abbreviations see main text.

exactly how to conduct each step of the process and how the results are described. Variation in methodology can affect results, and standardisation is required to keep up to date with frequent changes in methodology so that time series used for monitoring and assessment are comparable. Ideally, biomolecular protocols are extensively tested and intercalibrated through international collaborative projects such as MetaZooGene Intercalibration Experiment (Blanco-Bercial et al., 2025), BioGeoSCAPES for metaproteomic intercomparison, DNAqua-Net COST Action (Leese et al., 2016), and Horizon Europe projects BiOcean5D, OBAMA-NEXT and eDNAqua-Plan (Kavakiotis et al., 2026). Once standards are defined, they should be maintained and further developed by an authoritative body such as the International eDNA Standardisation Task Force (iESTF; see Box 5.1.) or the European Committee for Standardisation[138].

The development of an official standard is often a slow process, particularly in rapidly evolving fields such as biomolecular monitoring. Standardisation is an iterative process, fed by scientific advances. Initial procedures and definitions are shared, criticised, tested, and refined until they become functional, agreed, and verifiable. It is therefore useful to have documents that provide informal guidance while final standards are being developed. These documents may, for example, outline the best available methods or explain what information should be reported. These help the scientific and regulatory communities move in the same direction until there is sufficient consensus to develop an official standard. To speed up this process, several networks and repositories are available to support the marine biomolecular data life cycle. Under UNESCO's Ocean Best Practices System, Better Biomolecular Ocean Practices[139] has established protocol repositories that adhere to the Minimum Information about an Omics Protocol[140] standard. This initiative has also created a common global framework for data interoperability that aims to make -omics data machine-readable, and to develop standards to enhance data sharing.

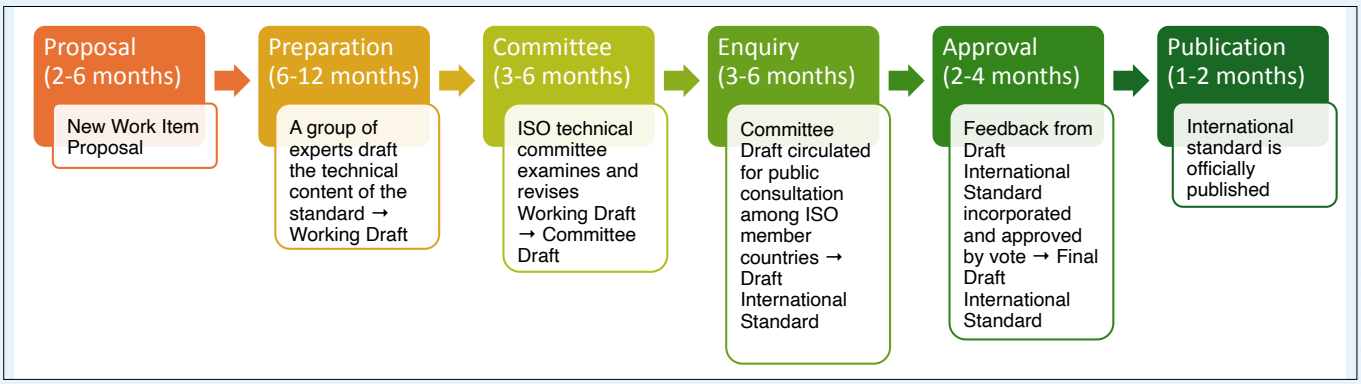
Although international standards have been slow to develop in biomolecular monitoring, all the components are in place for an efficient workflow towards standardisation (Figure 5.1.). Methodological foundations, baseline calibration and technical schemes have been provided by international consortia, while observation networks such as the Ocean Biomolecular Observing Network[141] (OBON) and the BioGeoSCAPES research programme are developing the capacity to deploy standardised protocols across Ocean basins, although additional resources are required. The best-established element of the workflow is the International Nucleotide Sequence Database Collaboration (INSDC)[142], which synchronises the major nucleotide repositories such as the European Nucleotide Archive (ENA)[143].

By enforcing metadata rules from best practice, standard operating procedures and formal standards, these digital archives can ensure that marine data remain searchable, comparable, and publicly accessible for global conservation efforts. Data flow from bioinformatic hubs (e.g. MGnify, Ocean-M[144]) is crucial for the implementation of biomolecular monitoring.

The Minimum Information about any (x) Sequence (MIxS) checklist[145], maintained by the Genomic Standards Consortium, specifies information that must be recorded for marine environmental samples. For marine -omics, mandatory information includes sampling context (i.e. geographic coordinates, depth, collection date and time, collection method), environmental variables (i.e. temperature, salinity, dissolved oxygen, pH, chlorophyll concentration), and analytical metadata (i.e. DNA extraction method and kit, PCR

138 http://www.cen.eu
139 https://bebop-obon.github.io/
140 https://github.com/BeBOP-OBON/miop
141 https://obon-ocean.org/
142 https://www.insdc.org
143 https://www.ebi.ac.uk/ena/browser/home
144 https://om.biocloud.net/home
145 https://genomicsstandardsconsortium.github.io/mixs/

Box 5.1. The International eDNA Standardisation Task Force (iESTF)



Workflow for the development of an ISO standard.

The iESTF[146] is a global initiative to develop and promote consistent, high-quality standards for eDNA methods in biodiversity monitoring and environmental assessments. It is composed of representatives from various disciplines including molecular biologists and environmental scientists, and other stakeholder types. Working groups tackle topics that are sufficiently mature to support international standards through the International Organisation for Standardisation (ISO, see figure). Currently the iESTF is developing several standards, including macroinvertebrate preservation for DNA-based analyses, extraction of aquatic environmental DNA from filters, metabarcoding to survey biological communities, and target species detection via quantitative PCR (qPCR)/digital PCR (dPCR). Although the pace of development of standards is insufficient for monitoring frameworks such as the WFD and MSFD), iESTF is currently facilitating and coordinating the development of precursor standards for ISO and other standards within the academic community. It also recognises the importance of the Global South in biodiversity assessments, monitoring and conservation. Therefore, it endorses standards that enforce quality assurance and quality control that are achievable with resource-efficient equipment. This creates equitable scientific standardisation and ensures that biomolecular tools can be accessible globally and for citizen science, while being robust and comparable.

primers, sequencing platform, read length, quality filtering). Adopting MIxS standards ensures datasets from different projects can be combined without extensive project-specific adjustments. For example, when researchers query the ENA for samples from 2020 to 2029 in the Northern European Seas, they find consistent documentation of depth, temperature, and sequencing protocols. Darwin Core[147], maintained by Biodiversity Information Standards[148] (TDWG) and used by the Ocean Biodiversity Information System[149] (OBIS) and the Global Biodiversity Information Facility[150] (GBIF), provides a vocabulary for describing species occurrences and sampling events, enabling linking biomolecular data with oceanographic measurements and traditional biodiversity surveys.

Funding agencies should promote participation in international networks like the OBON, the Group on Earth Observations Biodiversity Observation Network (GeoBON) or BioGeoSCAPES, and request TDWG-compliant data submissions. Journals should adopt minimum information standards for protocols and require DOI citations for methods. Repositories need to verify metadata before publication and link sequences directly to Darwin-Core compliant occurrence records. Regional biodiversity monitoring networks should formalise agreements, hold annual standards meetings, and develop capacity-building programs for emerging regions. Connections between regional networks require clear interoperability standards. Knowledge sharing would be facilitated if OBON was recognised as an official UN coordinating body, with regional steering committees, annual meetings, and a dedicated secretariat. Regular capacity assessments and customised support for developing regions would help ensure equitable standardisation advances. As eDNA monitoring becomes more important in marine conservation, standardisation infrastructure will become as vital as biomolecular techniques, turning isolated datasets into collaborative global knowledge networks.

146 https://iestf.global/
147 https://dwc.tdwg.org/
148 https://www.tdwg.org/
149 https://obis.org/
150 https://www.gbif.org/

5.4. FAIR and open data

For regional or global meta-analyses, data and metadata from individual studies should be easy to find, easy to access and be in a consistent format. Machine-readable, machine-interoperable[151], semantically structured[152], and provenance-rich data and metadata are key. These need to comply with the FAIR principles (Berry et al., 2021) and should be open. The FAIR data principles facilitate provenance tracking and comprehensive validation throughout the entire data life cycle. However, compliance with FAIR standards is not enough. Effective FAIR practice embeds robust quality control during sampling design and execution, DNA extraction, sequencing, and bioinformatics processing. It maintains traceable provenance from the moment a sample is collected in the field through all analyses and includes transparent treatment of uncertainty, including error estimates and explicit documentation of standard operating procedures (Santi et al., 2021). EMO BON (see Box 5.2.) base their data governance on FAIR compliance, and EMO BON and eDNAqua-plan exemplify European efforts to support the broader international marine -omics community with effective FAIR data practices in order to transform marine -omics research into credible, actionable knowledge that informs policy implementation (see Section 2.6.), that is integrated into the European and Global Ocean Observing Systems (EOOS and GOOS).

Once samples are collected and preserved, they must be accurately and consistently documented. Sample metadata should be submitted to international repositories (e.g. BioSamples[153]), where a unique identifier (ID) is assigned. This ID is a persistent unique identifier and is machine-actionable, linking physical samples to all subsequent derivatives, including extracted DNA, sequencing runs, analytical outputs, environmental data, and legal documents such as benefit-sharing agreements. Once a sample receives its BioSamples ID and is linked to the INSDC (which includes the ENA, the National Centre for Biotechnology Information (NCBI) GenBank[154], and the DNA Data Bank of Japan (DDBJ[155]), all subsequent records derived from that sample retain the same ID. Data models with keystone IDs allow for distributed curation, and corrections to automatically propagate to all downstream databases once records are updated.

Marine -omics produces Digital Sequence Information (DSI) from genetic resources and is therefore subject to national regulations that specify the boundary conditions of its usage for research, industrial- or policy-related applications. The Nagoya Protocol on Access and Benefit-Sharing (ABS), emerging international rules on DSI, and the BBNJ Agreement all mandate fair benefit-sharing with provider countries. For example, the COP16 Decision 16/2 (CBD, 2024), establishes a multilateral benefit-sharing mechanism for DSIs[156]. It urges commercial users to share benefits while exempting public databases, research, and academic institutions from fees. Benefit-sharing should encompass both financial and non-financial benefits, with all documentation linked to the datasets, including traceability from the original sampling operation. Marine -omics observatories must treat derived DSI in accordance with their geographic location, and permits may be required for some parties to cover its use. The European Marine Biological Resource Centre[157] (EMBRC) leads by example, ensuring compliance via centrally managing permits, agreements, and documentation for EMO BON at EMBRC-affiliated sites in Europe.

Equitable research practices will reduce global inequalities in scientific knowledge production. Responsible data management must therefore not only comply with the ABS clearing-house mechanism of the Nagoya Protocol, which provides information on ownership of genetic resources and conditions of use, but also considers the Collective Benefit, Authority to Control, Responsibility, Ethics (CARE) principles that emphasise equity and community rights. In other words, it defines the conditions under which the data can be accessed and used.

5.5. Data processing and workflow documentation

Data processing involves transforming biomolecular measurements (e.g. DNA sequences) into quality-controlled, analysable products. This significantly influences downstream interpretation, but remains one of the least standardised steps in marine -omics analysis. Uncertainties and errors may stem from low-quality sequencing, contamination or errors in reference databases. Different workflows may make fundamentally different choices: some filter aggressively and retain little data, while others contain errors. The trade-off between sensitivity (e.g. the ability to detect rare organisms) and specificity (e.g. the ability to avoid false detections) may depend on the research or management question. Workflows must therefore be carefully documented and follow the FAIR principles: processing steps should be clearly documented in machine-readable form using standardised workflow languages. Workflows should be shared through platforms like GitHub[158] or WorkflowHub[159]. Raw data should be deposited alongside versioned workflow documentation so future

[151] Machine-interoperable data can be read, exchanged and processed by different computer systems and software applications.
[152] Semantically structured data are organised with clear labels that explain exactly what the data mean and how they connect to other data.
[153] https://www.ebi.ac.uk/biosamples/
[154] https://www.ncbi.nlm.nih.gov/genbank/
[155] https://www.ddbj.nig.ac.jp/index-e.html
[156] https://www.cbd.int/dsi-gr
[157] https://www.embrc.eu/
[158] https://github.com/
[159] https://workflowhub.eu/

analysts can determine which choices were made and how quality-controlled sequences were derived.

ELIXIR, Europe's life science data infrastructure, shares services on European computing resources, and provides a stable environment for reference pipelines. MGnify, operated by the European Molecular Biology Laboratory-European Bioinformatics Institute[160] (EMBL-EBI), is an ELIXIR Core Data Resource, and serves as the framework for the EMO BON EMBRC MetaGOflow[161], enabling harmonised processing of marine metagenomic data from multiple observatories using community-endorsed methods and leveraging European high-performance and cloud computing. These infrastructures promote compatibility across data platforms and modalities, linking sequence datasets with imaging resources such as EcoTaxa[162], which hosts, annotates, and analyses plankton images and derives ecological data.

Box 5.2. The European Marine Omics Biodiversity Observation Network (EMO BON)



Credit: © Hanneloor Heynderickx

Biodiversity is sampled for DNA analysis from sessile organisms that settle on the hard substrate of Autonomous Reef Monitoring Structures (ARMS) as part of EMO BON.

Founded by EMBRC in 2021, EMO BON (Santi et al., 2023) is Europe's first centrally coordinated, long-term genomics observatory network spanning 22 coastal observatories from the Arctic to the Red Sea across 10 countries, creating a coordinated European Ocean -omics observation system. It applies standard protocols and infrastructure to integrate molecular biodiversity data into global efforts such as the UN Ocean Decade's OBON Programme. EMO BON samples across marine ecoregions, enabling comparisons of community structure across biogeographic areas. It covers the water column, sediments, and hard substrates (ARMS), and provides comprehensive biodiversity data from microorganisms to benthic communities. Standard protocols and metadata ensure comparability, transforming isolated efforts into a unified monitoring network. Repeated sampling includes bimonthly water-column measurements at all observatories, and soft-sediment benthic collections and annual deployments on hard substrates in some observatories, producing data on seasonal and interannual community change. By 2026, the network had generated thousands of DNA samples (currently 4,761 in total, with 1,991 sequenced and 2,906 stored in biobanks[163]), totalling around nine terabytes of genomic data, which support analyses of biodiversity trends. Time-series analyses reveal community variations by season, ecoregion, and latitude, detecting ecosystem responses to environmental and climatic variability. EMO BON provides early warnings of invasive species, harmful algal blooms, pathogens, and regime shifts, using spatially replicated genomic data linked to environmental indicators. By following FAIR principles and depositing data in public repositories such as ENA, it feeds into the European DTO. EMO BON is a scalable, standardised genomic observatory that supports EU policies, including the MSFD, ecosystem-based management, Marine Protected Areas, and environmental impact assessments amid climate and biodiversity crises.

[160] https://www.ebi.ac.uk/
[161] https://github.com/emo-bon/MetaGOflow
[162] https://ecotaxa.obs-vlfr.fr/
[163] A biobank is a repository that stores biological specimens for use in research.

5.6. The aggregation paradox: Quality over quantity

Large, data-rich studies may yield poorer results than small-scale studies. This counterintuitive effect, known as the aggregation paradox, occurs when variation from methodological differences dilutes ecological signals in the data. The causes may be differences in sampling methods (e.g. monthly versus quarterly sampling, different sampling volumes, and different collection depths), reference databases which are continually updated (i.e. with sequences being reassigned to different taxonomic identification over time), reporting biases (e.g. some projects publish all sequences whereas others filter out rare reads), and methodological errors (e.g. contamination that introduces false species occurrences that are not detectable in large-scale meta-analyses).

In a typical small-scale study, approximately 150-200 quality-screened species occurrences are published, with detection metadata, confidence scores, and links for review. This human-driven curation filters for higher signal-to-noise ratio thresholds, ensuring that marine managers trust positive detections as accurate indicators of presence. The aggregation paradox emphasises that quality outweighs quantity: a dataset with 200 verified, high-confidence species occurrences linked to sequences is more valuable than numerous unverified detections. Proper aggregation, standardisation via FAIR principles, and integration into federated systems such as DTO-BioFlow[164], GBIF and OBIS facilitate the application of this principle at scale.

Moving from raw eDNA sequences to shareable knowledge involves more than standardisation. It requires integrated ecological observatories that feed standardised, quality-validated data into federated systems. As eDNA monitoring becomes central for European water agencies and conservation programmes, infrastructure supporting standardisation and knowledge sharing will become vital. Projects such as the EU-funded Horizon Europe eDNAqua-Plan that invest in detailed metadata, adherence to standards, and publication of FAIR data will strengthen the shared knowledge resource, transforming ecological data from isolated research into collaborative knowledge systems.

5.7. Data products and multi-domain integration

Marine -omics must be inherently adaptable to scientific innovation. New sequencing technologies emerge, sampling networks expand or funding cycles and policy priorities expand. Static, inflexible data schemes hinder growth and innovation.

Many policy decisions require integrated interpretations from meta-analyses of multiple datasets across large geographic realms or long time periods, such as detecting the effect of global warming on species distribution and ecosystem functioning or tracing and managing the spread of invasive species. Using big data analytics within such federated knowledge frameworks requires rich data models and common standards. FAIR and open data generated according to standardised protocols can be fed into data aggregators (such as OBIS or the European Marine Observation and Data Network, EMODnet[165]) from where it can be sourced by federated, cloud-based platforms and fed into analytic infrastructures such as DTOs to provide answers to pressing marine policy issues. The EU Horizon European DTO-BioFlow project exemplifies this approach by establishing sustainable pipelines to integrate heterogeneous biodiversity data, including those derived from eDNA and metagenomic analyses, into a federated DTO to support science, policy and industry.

Effective FAIR data architectures use flexible, versioned schemes that allow the addition of new variables and relationships without affecting existing data. Different data modelling and structuring methods provide this flexibility. Darwin Core Archives[166] offer a standard format for biodiversity data that can be easily imported into global biodiversity platforms like GBIF and OBIS, and is particularly suitable for occurrence-based data. Linked data formats (e.g. JSON-LD[167]) enables self-describing records in which new fields can be added without affecting existing systems, while remaining structured enough for linking.

Resource Description Framework (RDF) triples[168] are useful for encoding complex relationships e.g. “Species A preyed-upon-by Species B in the Mediterranean in Summer 2024”, and automated queries enabling the discovery of ecological patterns. Darwin Core organises the basic facts, while a semantic layer adds precise meaning to those facts. Formats like JSON-LD and RDF connect everything together so that researchers can easily search and analyse massive amounts of complex information. An example is the EU Digital Europe Programme MareGraph[169] project.

To integrate biomolecular data into EOOS and GOOS would require embedding molecular sampling within broader Ocean observing platforms (see Section 4.2.). Autonomous eDNA samplers deployed on moored buoys or cabled observatories will be required to collect water samples at regular intervals alongside oceanographic sensor data. All sample metadata would need to be harmonised according to MIxS and Darwin Core standards and automatically submitted to EMBL BioSamples repository with links to raw sequences deposited at ENA. Derived biodiversity data need to be fed into the OBIS

[164] https://dto-bioflow.eu/
[165] https://emodnet.ec.europa.eu/en
[166] https://ipt.gbif.org/manual/en/ipt/latest/dwca-guide
[167] https://json-ld.org/
[168] RDF triples are a standard format for representing data and relationships on the internet and in semantic databases..
[169] https://www.maregraph.eu/

and GOOS BioEco Metadata Portal[170] to automatically integrate biodiversity data with physical and chemical oceanographic data. This would allow marine management agencies to query unified data to answer policy questions such as: "Which species shifted in response to this year's heatwave?", "Where did invasive species first appear?" and "What is the impact of large-scale climate change on fish stocks?". This integration will transform eDNA from a specialised research tool into an operational infrastructure that directly supports EU and global policy priorities for large-scale, cost-effective ecosystem monitoring.

5.8. Recommendations

Recommendations for research and infrastructure funders:

(5.1) Fund the creation, management and sustainability of standards through scientific discourse and intercalibration projects. Develop roadmaps and support the formation of specialist interdisciplinary standardisation task forces that include policymakers, policy implementers and enabling scientists across biological, chemical and physical sciences;

(5.2) Embed FAIR and CARE principles throughout grant agreements and request all data, metadata, protocols, and software to be as open as possible and as restricted as necessary;

(5.3) Establish technical infrastructure to manage FAIR open data (restricted when necessary), embedding Nagoya Protocol and BBNJ Agreement access and benefit-sharing frameworks directly into public database architectures;

(5.4) Invest in federated and quality-controlled infrastructure that ensures the sustainability of data generation and sharing on appropriate temporal and spatial scales for environmental monitoring; and

(5.5) Invest in tools and training for students and practitioners in order to render data, metadata, protocols, and software open and FAIR to ensure that data and metadata are discoverable through established technical standards and to ensure machine-actionability.

Recommendations for policymakers:

(5.6) Promote the adoption of FAIR and CARE principles for data collected in the framework of monitoring programmes and publish them as open as possible and as restricted as necessary;

(5.7) Commission solutions for emergent environmental issues from data-driven approaches including big data analyses and large-scale models; and

(5.8) Adopt management practices that are backed up with large and interdisciplinary datasets.



Credit: Masayuki, Ocean Image Bank.

Integrating biomolecular data into EOOS and GOOS would allow marine management agencies to query unified data to answer policy questions such as: "What is the impact of large-scale climate change on fish stocks?".

[170] https://bioeco.goosocean.org/

6 Capacity building and Ocean literacy

6.1. Summary

The establishment of biomolecular monitoring frameworks under European instruments such as the MSFD and the WFD requires investment in people and infrastructure. Capacity building supports the development of knowledge and competences that allow wider stakeholders to use biomolecular techniques and information. In addition, Ocean literacy ensures understanding of the Ocean's influence on society and society's influence on the Ocean. As such, the use of biomolecular tools for policy applications and environmental monitoring needs to be understood beyond specialist communities. This will help ensure that science-based marine policies are transparent and accessible to wider society, thereby supporting informed public dialogue and participation.

6.2. Capacity building

Capacity building initiatives aimed at the research community are relatively well developed. Organisations and networks such as EMO BON, EMBRC, the Partnership for Observation of the Global Ocean[171] (POGO), EuroMarine[172], and several Horizon Europe projects support a range of specialist training opportunities in marine molecular biology. These contribute to building knowledge and capacity in aquatic molecular biodiversity and include: the Marine Training Unit[173] (MTU) within EMBRC; the Ocean Training Partnership[174] of POGO; OBON Working Groups[175]; GuardIAS eDNA Summer School[176]; and stakeholder workshops organised through projects such as eDNAqua-Plan. Short Term Scientific Missions under European Cooperation in Science and Technology (COST) Actions[177] and EU Infrastructure programmes allow researchers access to specialist expertise in marine genomics, biodiversity assessment, bioinformatics, and molecular data interpretation (e.g. AquaINFRA[178], DNAqua-Net), as well as to research vessels, marine stations, laboratory facilities, sequencing platforms, and computational resources (e.g. ASSEMBLE+[179], Eurofleets[180]) that may not be available within their home institutions or countries. Advanced training programmes such as those offered by the EMBL[181] and the Emerging Bioinformatics Approaches for Microbial Ecogenomics (EBAME) workshop series[182] aim to strengthen data literacy and interpretative skills, equipping scientists to extract maximum value from biomolecular information. On-demand training platforms, such as EMBL-EBI Training[183], allow researchers and practitioners to build these skills asynchronously, without the need to travel or to leave their workplace. Equitable access to participants in the Global South to these advanced training opportunities is needed to support capacity building globally, which is key to support the implementation of the BBNJ Agreement. This requires the provision of scholarships and grants to support participation in training programmes, workshops, and other capacity-building activities, particularly for representatives of Small Island Developing States (SIDS).

An additional challenge relates to scale. For example, the Atlantic Meridional Transect -omics network[184] (AMT-omics) offers participation in research expeditions and learning opportunities through POGO fellowships, however these are limited to a few participants. In contrast, collaborations such as the European University of the Seas[185] (SEA-EU) are able to develop accredited training in marine sciences at a potentially larger scale. Private companies can contribute to capacity building by supporting skill development and learning on their technology platforms. For example, Oxford Nanopore Technologies offer online training programmes designed to support use of their sequencing technologies. Their new portable sequencing technology, combined with compact molecular laboratory platforms (i.e. mobile laboratories) such as those developed by Bento Lab[186], enables the deployment of advanced biomolecular tools in virtually any location, including aboard small vessels and in remote or harsh coastal environments. This approach enhances access to molecular analyses while promoting laboratory training, capacity building, and knowledge exchange on a global scale.

171 https://pogo-ocean.org/
172 https://euromarinenetwork.eu/
173 https://www.embrc.eu/marine-training/
174 https://oceantrainingpartnership.org/
175 https://obon-ocean.org/about/working-groups/
176 https://www.marinetraining.eu/node/6251
177 https://www.cost.eu/
178 https://aquainfra.eu/
179 https://www.vliz.be/projects/assembleplus/home.html
180 https://www.eurofleets.eu/
181 https://www.embl.org/about/info/course-and-conference-office/events/mmd26-01/
182 https://maignienlab.gitlab.io/ebame/
183 https://www.ebi.ac.uk/training/
184 https://obon-ocean.org/project/amt-omics-atlantic-meridional-transect-omics-network/
185 https://sea-eu.org/
186 https://bento.bio/bento-lab/portable-pcr-for-nanopore/

Outside the research community, opportunities to acquire biomolecular competencies are less well developed, particularly for marine managers or practitioners who commission biomolecular monitoring and rely on biomolecular outputs without directly implementing the methods themselves. At present, very few training and awareness initiatives are aimed directly at policy and management contexts. One example is the GuardIAS eDNA Summer School organised by Nord University which aimed to raise awareness of the benefits of eDNA tools for the early detection of aquatic alien species and their integration into policy and management frameworks. The programme was open to "professionals and practitioners", but the core cohort consisted of PhD students from universities across Europe. This traditional training model, which requires participants to spend an entire week away from their workplace, is less suitable for stakeholders and policymakers, who often have other unavoidable commitments. EU COST Actions such as DNAquaNet and eDNA4Policy are bridging the science–policy gap by linking researchers and authorities to co-develop environmental monitoring protocols. However, widespread training will be necessary to fully equip marine managers, environmental practitioners, and policymakers with the knowledge and understanding required to interpret and apply biomolecular evidence within monitoring and management frameworks. Future COST or Marie Skłodowska-Curie Actions[187], together with other European funding initiatives, should focus on training solutions adapted to their needs. Alternatively, networks such as EMBRC and SEA-EU, have the expertise to provide training resources, and could offer micro-credentials to help formalise progression and support scalable capacity building. In addition, the EU Horizon Europe ANERIS project[188] is developing genomic tools to improve the monitoring and management of marine ecosystems using a co-design framework that engages and trains stakeholders from academia, industry, government, and civil society to develop and use data products. Massive Open Online Courses provide free or low-cost access to courses aimed at all levels from introductions to eDNA sampling through to courses on advanced bioinformatics. However, administrative barriers need to be minimal to avoid any unnecessary barriers to knowledge sharing.

Making training resources free and accessible can help support equitable participation and fair scientific practice beyond the EU by allowing professionals in all geographic areas to upskill where required and may help to limit persistent issues such as "parachute research", where data are collected in less-resourced regions without meaningful involvement of local scientific communities (Sarmento et al., 2025). Regular capacity assessments should be conducted and customised support should be provided to developing regions to help ensure that regional expertise in bioinformatics and data science advances equity. All training globally needs to emphasise the importance of standardisation, open science, and FAIR principles to further advance equity.

Further challenges relate to the clarity and usability of training material, which may be further compounded by language barriers and the limited availability of translated information. The rapid evolution of biomolecular techniques creates a continuous need for updates, particularly with respect to data interpretation and bioinformatics literacy. Consequently, educational resources require frequent updates to remain accurate and relevant for both laboratory practice and data interpretation. Blended learning approaches that combine theoretical instruction, visual resources, and supervised practical experience is an effective strategy for ensuring that educational content remains up to date while promoting the development of laboratory competencies in molecular biology (Sarmento et al., 2025).

6.3. Ocean literacy

Biomolecular techniques are often perceived as complex or abstract. Explaining concepts such as eDNA or -omics in simple terms helps people understand the relevance of these tools and supports their acceptance in marine policy and management. Outreach materials that simplify concepts are key, as well as curriculum development, teacher training, and open-access educational resources for diverse audiences. For example, the Tara Ocean Foundation[189] has combined large-scale scientific exploration with strong storytelling, visual communication and educational outreach, helping to make complex biomolecular research visible and understandable to wider audiences. The Traversing European Coastlines (TREC) expedition (see Section 4.5.) similarly paired large-scale coastal sampling with a dedicated public engagement and Ocean literacy programme. Non-governmental organisations (NGOs) also play an important role in Ocean literacy. For example, publications such as the IUCN (2024) "Issues Brief" on eDNA[190] can reach a wider civil society audience.

An important aspect of Ocean literacy is citizen science initiatives that incorporate biomolecular activities, including eDNA sampling, which allow non-expert participants to engage with monitoring efforts in an accessible way. Global initiatives such as UNESCO's eDNA Expeditions[191], or the former Ocean Sampling Day[192], illustrate how schools, NGOs and local communities can participate in biomolecular observation through standardised, easy-to-use sampling kits, while laboratory analyses and interpretation remain expert-led. By involving non-experts in appropriate and well-designed monitoring activities, citizen science initiatives cultivate shared responsibility and a sense of collective stewardship for the marine environment (see EMB Position Paper N°. 23 on Advancing Citizen Science for Coastal and Ocean Research; Garcia Soto et al., (2017). The Danish HavBlitz project demonstrates how volunteers can contribute to large-scale biodiversity assessments using simplified eDNA sampling linked to professional analysis (see Box 6.1.).

187 https://marie-sklodowska-curie-actions.ec.europa.eu/
188 https://aneris.eu/
189 https://fondationtaraocean.org/en/foundation/
190 https://iucn.org/resources/issues-brief/environmental-dna
191 https://www.unesco.org/en/edna-expedition
192 https://www.microb3.eu/osd.html

Box 6.1. HavBlitz Ocean literacy and citizen science initiative, Denmark



Credit: Niels-Ole Hørlyk, Agersnap et al. (2022) (CC BY 4.0).

The HavBlitz initiative in Denmark involved the public in the observation of marine biodiversity while increasing awareness and understanding of biomolecular approaches.

HavBlitz was a national campaign in Denmark to involve the public in the observation of marine biodiversity while increasing awareness and understanding of biomolecular approaches (Agersnap et al., 2022). It combined simplified eDNA sampling with participatory engagement activities along coastal and marine areas. HavBlitz was implemented as a coordinated sampling campaign with data collection activities conducted within a limited seasonal window. This allowed broad geographic coverage and comparability of samples collected by participants across multiple locations. Volunteers collected seawater samples from approximately 100 coastal and nearshore sites using small, simple and standardised *in situ* filtration kits. DNA extraction, sequencing and metabarcoding analysis were subsequently conducted by research institutions using established laboratory workflows. By participating in a structured and time-limited monitoring campaign, volunteers gained first-hand experience of core concepts such as eDNA, biodiversity assessment and marine ecosystem observation. The metabarcoding data generated during the campaign enabled a comprehensive assessment of fish biodiversity across a broad spatial scale, demonstrating the potential of eDNA-based approaches for marine monitoring. The initiative provided insight into how biomolecular tools can contribute to large-scale monitoring and management, without needing to perform complex laboratory or analytical tasks.

6.4. Recommendations

Recommendations for research funders:

(6.1) Require mandatory capacity-building, training, and stakeholder engagement activities within research projects to ensure that biomolecular results are translated into practical management knowledge and can be effectively used by managers, practitioners, and policymakers; and

(6.2) Foster long-term citizen science initiatives involving simplified biomolecular approaches with expert support provided to ensure robust data analysis and quality assurance, thereby contributing to the establishment of biodiversity baselines, and wider societal acceptance of biomolecular methods.

Recommendations for scientists:

(6.3) Develop and maintain freely accessible and universal biomolecular training material, online courses and multilingual resources for students through international research networks; and

(6.4) Develop and maintain accessible training resources for marine managers, practitioners, citizens and policymakers (e.g. online modules, micro-credentials and professional development courses), together with plain-language outreach and educational materials, including curriculum development, teacher training and open-access resources, and awareness programmes.

Recommendations for policymakers:

(6.5) Conduct regular capacity assessments and provide sustained support to developing regions, including access to laboratory and computational resources, training and technical support to ensure that expertise advances equitably.

7 Overarching recommendations

This document demonstrates the potential of biomolecular approaches to revolutionise the observation and management of marine life. The vision presented in this document of a global -omics observatory, including innovation in low-cost *in situ* sensors and samplers and artificial intelligence, can advance the capacity to monitor and respond to changes in marine biodiversity and expedite the flow of *in situ* data to policymakers for adaptive management. The development of policy indicators derived from biomolecular data can make monitoring more efficient, accurate and cost effective and transform the implementation of European environmental policies, as well as help meet international targets for biodiversity in relation to the Kunming-Montreal Global Biodiversity Framework and the BBNJ Agreement.

However, the use of biomolecular methods for monitoring, managing and protecting marine environments needs to be scaled up to realise its potential to support policy objectives. Specific recommendations that will facilitate and accelerate the incorporation of biomolecular approaches into the observation and management of marine life are included at the end of each chapter of this document and summarised in Table 7.1., where they are organised into categories of: policy integration; fundamental biomolecular science; technological innovation; hybrid monitoring and field calibration; data and digital infrastructure; global standards and ethical governance; and capacity building and global networks.

Recommendations for research funders:

- **Fund co-designed research, including the exploitation of large, interdisciplinary datasets that link biomolecular methods to policy applications.** This includes research on molecular traits, -omics functional knowledge and plankton metabolism. Machine-readable indicators will enable data integration into climate models and Digital Twins of the Ocean to improve future spatial climate scenario projections;
- **Fund exploratory and curiosity driven fundamental science to investigate baseline molecular mechanisms** in marine organisms and Ocean ecosystems to advance understanding of ecosystem function and dynamics. Research is needed to: characterise genes with unknown functions; determine the speed of DNA degradation; analyse existing datasets; advance multi-omics; better link organisms, genes and functions; and improve ecological models;
- **Invest in accessible and interoperable computational and digital infrastructure,** including repositories, visualisation tools and federated reference libraries. Support is needed for data generation, storage, curation, quality control and adherence to FAIR metadata standards, while enabling long-term monitoring and compliance with the Nagoya Protocol and the BBNJ Agreement. This will facilitate integration of biomolecular data into Digital Twins of the Ocean for EU policy implementation;
- **Fund the targeted expansion and improvement of molecular reference databases through sequencing campaigns** that generate barcodes, genomes, transcriptomes and proteomes from organisms that are currently not, or poorly, represented in reference databases;
- **Support long-term -omics observatories alongside traditional sampling and monitoring networks (i.e. augmented observatories)** that feed into holistic ecosystem monitoring and assessments;
- **Fund innovation in automated and miniaturised *in situ* biomolecular samplers and sensors,** including scaling-up and commercialisation of existing technologies, and the expansion of AI for rapid sequence identification and association of -omics data with other environmental parameters to predict planetary-scale trends;
- **Enforce rigorous standardisation for biomolecular protocols and ethical open-access frameworks** across the entire -omics workflow to ensure data reuse and interoperability. Embed FAIR and CARE principles in research funding;
- **Fund interdisciplinary biomolecular standardisation task forces** within international research networks to ensure that biomolecular discoveries translate into actionable management knowledge; and
- **Fund sustained training, collaboration and stakeholder engagement** connecting technology developers, data experts and practitioners to improve data and metadata quality, develop AI-ready datasets, and support long-term citizen science. Provide sustained support and resources to developing regions to ensure expertise advances equitably.

Recommendations for policymakers:

- **Integrate insights on biodiversity and ecosystem function from biomolecular data into the design of protected areas and restoration initiatives** to support EU policies including the Habitats Directive and Nature Restoration Regulation;
- **Embrace novel frameworks to value ecosystem services based on molecular functions;**
- **Commission solutions for emergent environmental issues from data-driven approaches** including big data analyses and large-scale models, and adopt management practices that are backed by large and interdisciplinary datasets; and
- **Promote the adoption of FAIR and CARE principles for data collected within monitoring programmes** and the publication of data as open as possible and as restricted as necessary.



Credit: Liang Fu, Ocean Image Bank.

Biomolecular approaches can be used to determine species interactions as part of complex biological systems.

Table 7.1. Summary of recommendations for research funders, policymakers and scientists.

ACTION AREA	RECOMMENDATION	RESEARCH FUNDERS	POLICY-MAKERS	SCIENTISTS
Policy Integration	2.1: Fund co-designed research (involving policymakers, scientists and industry) linking biomolecular methods to policy applications, including co-development of new machine-readable policy indicators ready for integration into the European Digital Twin of the Ocean.	X		
	2.8: Develop a framework to value ecosystem services based on molecular functions.			X
	2.9: Integrate insights on biodiversity and ecosystem function from biomolecular data into the design of protected areas and restoration initiatives to support policies including the Habitats Directive and Nature Restoration Regulation.		X	
	2.10: Embrace novel frameworks to value ecosystem services based on molecular functions.		X	
	4.3. Fund research to incorporate plankton metabolism and molecular traits into Digital Twins of the Ocean and IPCC models to support spatial and climate scenario projections of -omics functional knowledge.	X		
	5.7. Commission solutions for emergent environmental issues from data-driven approaches including big data analyses and large-scale models.		X	
	5.8. Adopt management practices that are backed up with large and interdisciplinary datasets.		X	
Fundamental Biomolecular Science	2.4: Fund research on the fundamental understanding of biodiversity and ecosystem function using biomolecular tools to stimulate potential advances in policy-relevant areas, and with potential practical applications in biotechnology.	X		
	2.5. Fund projects on the novel interpretation of existing datasets to stimulate trans- and inter- disciplinary data analysis.	X		
	2.6. Fund proof-of-concept research on the use of biomolecular data to understand baseline ecological dynamics and the potential effects of industrial activities.	X		
	3.1. Fund research to investigate DNA degradation kinetics, including temporal stability under different environmental conditions, the distances over which DNA is transported, and its sinking rate.	X		
	3.5. Fund research to characterise genes of unknown function from existing and new model organisms to develop more reliable and complete reference databases of gene function.	X		
	4.6. Fund multi-omics research across diverse species to better link organisms to their functions, improve accuracy of models and spatial and temporal predictions, and explore new "-omics"-informed theories.	X		

ACTION AREA	RECOMMENDATION	RESEARCH FUNDERS	POLICY-MAKERS	SCIENTISTS
Technological Innovation	4.1. Fund innovation in the development of low-cost, miniaturised deployable samplers and sensor technologies in order to accelerate their use at scale.	X		
	4.2. Fund the expansion of the use of AI for faster sequence identification using well-curated databases, and to associate -omics data with other environmental parameters to predict planetary-scale trends.	X		
	4.4. Accelerate a European eDNA technology pipeline in the first instance, with other technologies to follow, by strengthening support for technological development and high-impact demonstrations, while incentivising commercialisation.	X		
Hybrid Monitoring and Field Callibration	2.2. Support the development of augmented observatories by including biomolecular data, alongside traditional observations, into longitudinal, meso-scale and process studies, allowing long-term calibration with traditional methods.	X		
	4.7. Perform intercomparison exercises with environment agency standards, to expand the adoption in the field of user-friendly portable technologies starting with underway autonomous and manually deployable eDNA sampling systems, saving costs and time at sea.			X
	4.8. Integrate information on biodiversity with other data sources (e.g. oceanographic data, remote sensing, taxonomy databases) in monitoring surveys for holistic marine ecosystem assessments.		X	
Data & Digital infrastructure	2.3. Invest in easy-to-use metadata and data repositories, and visualisation interfaces for molecular data, to ensure broad use of scaled-up molecular observations, data sharing and accessibility in an open science framework.	X		
	3.2. Fund the further development of the architecture for a federated system of reference databases that are independently curated and maintained. Promote adherence to the principles of metadata publishing, quality control, and taxonomic curation, connected within a common network enabling information exchange, interoperability, and findability as envisaged by eDNAqua-Plan.	X		
	3.3. Fund collection and sequencing campaigns to complete eukaryotic coverage for European waters for DNA barcode databases to facilitate metabarcoding and eDNA biodiversity analyses in support of MSFD, WFD, Nature Restoration Regulation and MPA monitoring.	X		
	3.4. Fund research to generate genomes, transcriptomes and proteomes from a large diversity of accurately identified species from taxonomic groups that are not yet well represented in databases. This is required to develop more reliable and complete reference databases and increase the value and interpretability of existing and future biomolecular datasets.	X		
	5.3. Establish technical infrastructure to manage FAIR open data (restricted when necessary), embedding Nagoya Protocol and BBNJ Agreement access and benefit-sharing frameworks directly into public database architectures.	X		
	5.4. Invest in federated and quality-controlled infrastructure that ensures the sustainability of data generation and sharing on appropriate temporal and spatial scales for environmental monitoring.	X		

ACTION AREA	RECOMMENDATION	RESEARCH FUNDERS	POLICY-MAKERS	SCIENTISTS
Global Standards & Ethical Governance	2.7. Establish a working group, potentially under the Group on Earth Observations Biodiversity Observation Network (GeoBON), on the cross-cutting implementation of biomolecular tools in all classes of Essential Biological Variables.			X
	5.1. Fund the creation, management and sustainability of standards through scientific discourse and intercalibration projects. Develop roadmaps and support the formation of specialist interdisciplinary standardisation task forces that include policymakers, policy implementers and enabling scientists across biological, chemical and physical sciences.	X		
	5.2. Embed FAIR and CARE principles throughout grant agreements and request all data, metadata, protocols, and software to be as open as possible and as restricted as necessary.	X		
	5.5. Invest in tools and training for students and practitioners in order to render data, metadata, protocols, and software open and FAIR to ensure that data and metadata are discoverable through established technical standards and to ensure machine-actionability.	X		
	5.6. Promote the adoption of FAIR and CARE principles for data collected in the framework of monitoring programmes and publish them as open as possible and as restricted as necessary.		X	
Capacity Building & Collaborative Networks	4.5 Provide dedicated funding to link technology developers, citizen scientists and data experts to standardise and align the description, structure and format of data and metadata to supply AI-ready data with controlled vocabularies and ensure reporting to shared repositories and dashboards that include advanced data visualisation tools.	X		
	6.1. Require mandatory capacity-building, training, and stakeholder engagement activities within research projects to ensure that biomolecular results are translated into practical management knowledge and can be effectively used by managers, practitioners, and policymakers.	X		
	6.2: Foster long-term citizen science initiatives involving simplified biomolecular approaches with expert support provided to ensure robust data analysis and quality assurance, thereby contributing to the establishment of biodiversity baselines, and wider societal acceptance of biomolecular methods.	X		
	6.3. Develop and maintain freely accessible and universal biomolecular training material, online courses and multilingual resources for students through international research networks.			X
	6.4. Develop and maintain accessible training resources for marine managers, practitioners, citizens and policymakers (e.g. online modules, micro-credentials and professional development courses), together with plain-language outreach and educational materials, including curriculum development, teacher training and open-access resources, and awareness programmes.			X
	6.5. Conduct regular capacity assessments and provide sustained support to developing regions, including access to laboratory and computational resources, training and technical support to ensure that expertise advances equitably.	X	X	

List of abbreviations and acronyms

12S	A gene found in the mitochondrial DNA of animals, plants, and other eukaryotes that codes for part of the mitochondrial ribosome
16S	A gene found in eukaryotes prokaryotes that codes for part of the ribosomal RNA small subunit
18S	A gene found in eukaryotes that codes for part of the ribosomal RNA small subunit
28S	A gene in eukaryotic cells that codes for the ribosomal RNA large subunit
ABS	Access and Benefit-Sharing
aDNA	Ancient DNA
AI	Artificial Intelligence
ARMS	Autonomous Reef Monitoring Structures
AUV	Autonomous Underwater Vehicle
BBNJ	Agreement on Marine Biological Diversity of Areas Beyond National Jurisdiction
BWD	Bathing Water Directive
CARE	Collective benefit, Authority to Control, Responsibility, Ethics
CBD	Convention on Biological Diversity
cdPCR	Chip-based digital PCR
CFP	Common Fisheries Policy
COI	Mitochondrial cytochrome oxidase subunit-I gene (a gene found in the mitochondrial DNA of most animals and encodes for the enzyme)
COP	Conference of the Parties
CRISPR-Cas	Clustered Regularly Interspaced Short Palindromic Repeats–CRISPR-associated proteins/ system
CTD	Conductivity, Temperature, Depth sensors
DDBJ	DNA Data Bank of Japan
ddPCR	Digital Droplet PCR
dPCR	Digital PCR
DNA	Deoxyribonucleic acid
DOC	Dissolved Organic Carbon
DOI	Digital Object Identifier
DSI	Digital Sequence Information
DTO	Digital Twin of the Ocean

EBV	Essential Biological Variable
eDNA	Environmental DNA
EIA	Environmental Impact Assessment
EMBRC	European Marine Biological Resource Centre
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
EMO BON	European Marine Omics Biodiversity Observation Network
EMODnet	European Marine Observation and Data Network
ENA	European Nucleotide Archive
EOOS	European Ocean Observing System
EOV	Essential Ocean Variable
ESP	Environmental Sample Processor
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GBIF	Global Biodiversity Information Facility
GEM	Genome-scale metabolic model
GeoBON	The Group on Earth Observations Biodiversity Observation Network
GES	Good Environmental Status
GOOS	Global Ocean Observing System
HAB	Harmful Algal Bloom
HPLC-MS	High-Performance Liquid Chromatography Mass Spectrometry
ID	Unique Identifier
IOC–UNESCO	Intergovernmental Oceanographic Commission of UNESCO
iESTF	International eDNA Standardization Task Force
INSDC	International Nucleotide Sequence Database Collaboration
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
IPCC	Intergovernmental Panel on Climate Change
ISO	The International Organization for Standardization
ITS	A region of non-coding DNA found in eukaryotic organisms
IUCN	International Union for Conservation of Nature

JSON-LD	JSON Linked Data
MALDI-TOF MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MBON	Marine Biodiversity Observation Network
mCDR	marine Carbon Dioxide Removal
MGR	Marine Genetic Resource
MIxS	Minimum Information about Any Sequence
miRNA	MicroRNA
ML	Machine Learning
MOTI	Marine Omics Technology and Instrumentation
MPA	Marine Protected Area
mRNA	messenger RNA
MSFD	Marine Strategy Framework Directive
MTU	Marine Training Unit
NCBI	The National Centre for Biotechnology Information
NGO	Non-governmental organisation
NGS	Next Generation Sequencing
NIS	Non-indigenous species
NMR	Nuclear Magnetic Resonance
NOAA	National Oceanic and Atmospheric Administration
OBIS	Ocean Biodiversity Information System
OBON	Ocean Biomolecular Observing Network
PCR	Polymerase Chain Reaction
POGO	Partnership for Observation of the Global Ocean
qPCR	Quantitative Polymerase Chain Reaction
rbcL	A gene that codes for the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit of chloroplast cells of plants and algae
RDF	Resource Description Framework
RNA	Ribonucleic acid
RNA-seq	RNA sequencing

ROV	Remotely Operated Vehicle
RPA	Recombinase Polymerase Amplification
rRNA	ribosomal RNA
SEA-EU	European University of the Seas
SIDS	Small Island Developing States
SNP	Single Nucleotide Polymorphism
TDWG	Biodiversity Information Standards, historically Taxonomic Databases Working Group
TREC	Traversing European Coastlines
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
USV	Unmanned Surface Vehicles
WFD	Water Framework Directive
WHO	World Health Organization

Glossary

Acetylation: The addition of an acetyl group to a molecule. In epigenetics, acetylation of histone proteins can influence how accessible DNA is for gene expression.

Amplicon Sequence Variants: Highly resolved identical sequences that allow precise identification of species.

Ancient environmental DNA or sedimentary ancient DNA (aDNA): Marine sedimentary ancient DNA (aDNA) is eDNA from organisms that lived in the ocean in the past and that has accumulated over time on the seafloor and is preserved in sediment.

Archaea: Single-celled prokaryotic microorganisms that are widespread in (extreme) marine and terrestrial environments.

Autosampler: A device that automatically collects water samples from the Ocean at set times or places.

Auxotrophy: The inability of an organism to synthesise a particular compound required for its growth and survival which it must therefore obtain from its environment.

Bacteriophage: A type of virus that infects and replicates inside bacteria.

Barcoding: A technique that involves the amplification and sequencing of genetic markers, or barcodes, that vary among species but are similar within a species. These are compared to reference databases for precise taxonomic assignment.

Biobank: A repository that stores biological specimens for use in research.

Bioinformatic products: Outputs or tools created by the computational analysis of biological data.

Bioinformatics: The use of computer science, mathematics and statistics to analyse large and complex biomolecular data.

Bioprospecting: Searching for, discovering, and utilising biological resources (such as genes, enzymes, or other molecules) from nature for commercial or scientific purposes.

Bioprovinces: Ocean regions with distinct environmental features and biological communities.

Biome: A large ecological region or system, such as the deep sea, characterised by broad environmental conditions and the organisms that live there.

Biosynthetic pathway: A series of chemical processes that build complex molecules.

Ciliate: A single-celled eukaryotic microorganism covered in tiny, hair-like structures called cilia.

Commensalism: A type of relationship between two different species in which one benefits while the other is neither helped nor harmed.

Community composition: Which species are present and in what abundance.

Compositional data: These represent relative parts of a whole. -Omics data are compositional, and for example, a percentage of reads is interpreted as a percentage of an organism within a community.

Constraint-based programming: Applying a set of rules to analyse and predict the behaviour of complex systems.

CRISPR-Cas: A tool to edit DNA with high precision, using a system from the natural defences of bacteria.

CRISPR-Cas9 gene knockout: A genetic technique that uses the CRISP-Cas9 system to disrupt a specific gene, allowing researchers to determine its function by observing the effects of removing or disabling it.

Cryptic species: Species that are morphologically indistinguishable from one another but are genetically distinct.

Deoxyribonucleic acid (DNA): The genetic hereditary material in cellular organisms and many viruses. It stores genetic information, including protein-coding and regulatory sequences.

Diatom: Microscopic, single-celled or colonial algae that have shells made of silica.

Digital Droplet PCR (ddPCR): A laboratory technique using PCR amplification to very precisely quantify the amount of a specific DNA marker in a sample.

Digital PCR (dPCR): A laboratory technique that amplifies and quantifies specific DNA segments.

DNA barcode: A fragment of an organism's genome that can be used to reliably distinguish and identify species.

Ecosystem function: Ecosystem function refers to the natural processes and interactions that take place within an ecosystem allowing it to sustain life e.g. nutrient cycling. Ecosystem services are dependent on a properly functioning ecosystem.

Edge computing: Localised data processing.

Endosymbiotic cyanobacterium: A type of microscopic, photosynthetic bacteria that lives inside another organism in a symbiotic relationship.

Environmental DNA (eDNA): The genetic material found in the environment (e.g. water, sediment, soil, air) shed from organisms (e.g. cells, faeces) that can be used to study biodiversity without needing to capture or directly see the organism itself.

eDNA Degradation: The breakdown of environmental DNA over time through microbial activity, temperature, ultraviolet radiation and other physical and chemical conditions. The rate of degradation determines how long an organism's genetic signal persists after it has left a location, and therefore how closely an eDNA detection reflects recent rather than past presence.

Epigenetics: Heritable chemical and structural changes of DNA and chromatin that regulate gene expression without altering the DNA sequence itself. In marine research, epigenetic markers are used to estimate the age of individual organisms, to track responses to stressors such as Ocean warming, acidification, deoxygenation and pollution, and as candidate early-warning biomarkers of environmental change.

Essential Biodiversity Variables (EBV): Global standardised, scientifically-defined indicators developed to track key aspects of biodiversity and used to guide reporting on biodiversity within major treaties.

Essential Ocean Variables (EOV): A minimum set of Ocean variables that are needed to assess Ocean state and variability for important global Ocean phenomena (see GOOS phenomena list), and to provide essential data for applications that support societal benefit.

Eukaryote: An organism whose cells contain a membrane-bound nucleus and other membrane-bound organelles, and includes animals, plants, and protists.

Explainable AI: A set of tools and methods that allows humans to understand and trust the outputs created by machine learning.

Faecal Indicator Bacteria: Microorganisms whose presence indicate faecal contamination of water.

Foraminifera: Single-celled marine eukaryotes, often with shells.

Free DNA: DNA existing outside of living cells.

Gametes: Reproductive cells.

Gene expression: A cellular process by which the instructions coded by DNA are converted into functional products, such as proteins. It can be investigated through transcriptomics.

Genomic mining: The process of searching through genomes to identify genes that may have interesting or useful functions.

Genomics: The study of DNA sequences, including genes, within a cell, tissue or organism (i.e. its genome) and their interactions with each other and the environment.

Genetic drift: The change in frequency of a gene variant in a population from generation to generation due to chance, rather than natural selection.

Harmful Algal Bloom (HAB): Overgrowth of algae in marine or freshwater systems that can produce toxins, deplete oxygen, and disrupt ecosystems, monitored using molecular and conventional methods.

Heterotrophic: Describes organisms that obtain energy and carbon by consuming organic matter produced by other organisms.

High-Performance Liquid Chromatography Mass Spectrometry (HPLC-MS): A technique to separate, identify and measure compounds in a complex mixture.

High-throughput molecular sequencing technologies: Technologies that allow sequencing of multiple DNA molecules simultaneously, rapidly creating large datasets.

Histones: Proteins found in the nucleus of cells that help organise and package DNA.

Holobiont: An ecological unit comprising a host organism and its associated microbial communities.

Integrative Taxonomy: The combined use of molecular and ecological data with traditional methods to define species boundaries with greater precision than morphology-based approaches alone.

Interactome: A complete set of interactions between molecules within a cell, organism or biological system.

Isomers: Molecules that have the same chemical formula but a different arrangement of atoms.

Lab on a chip: Miniaturised devices that integrate lab functions onto a small chip.

Lateral flow strips: Paper-based devices used to detect the presence of a specific substance.

Linnaean classification: The traditional system of naming organisms into a ranked hierarchy based on shared physical characteristics e.g. kingdom, phylum, class, order, family, genus, species.

Long-read technologies: Technologies involving sequencing outputs that are typically 10,000 - 100,000 base pairs.

Machine-interoperable data: Data that can be read, exchanged and processed by different computer systems and software applications.

MALDI-TOF MS: A technique to rapidly identify and analyse proteins by measuring their mass.

Mass spectrometry: A technique used to measure the mass of molecules and identify their chemical structure.

Meiofaunal: Small invertebrates living in aquatic sediments.

Metabarcoding: A technique that involves the PCR amplification and high-throughput sequencing of genetic marker regions from DNA extracted from mixed-community or environmental samples, enabling taxonomic characterisation of multiple organisms.

Metabolic cross-talk: Biochemical communication between organisms and cells which influences function, survival and growth.

Metabolic engines: A concept that considers each organism as biochemical machinery whose interaction contributes to biogeochemical cycles.

Metabolic processes: Sets of chemical reactions carried out by living organisms that sustain life by producing energy and building cellular components. Metabolic processes within cells or organisms convert one molecule into another, forming ordered and connected metabolic pathways.

Metabolites: Compounds such as amino acids, fatty acids, vitamins, toxins, and signalling molecules that are produced by cells during metabolic processes. These play critical roles in energy production, growth, defence, communication, and interactions between organisms.

Metabolomics: The study of the metabolites, or small molecules, present in an ecosystem, organism, cell or tissue under certain conditions.

Metagenome: The genetic material from the organisms present in an environmental sample.

Metagenomics: The study of genetic material obtained from organisms in an environmental sample (e.g. eDNA), enabling comprehensive identification of the organisms that are present, and characterisation of the collective genetic potential of the community.

Metatranscriptomics: The study of which genes are being actively expressed (e.g. from mRNA from organisms in an environmental sample), showing which metabolic pathways are active under specific environmental conditions.

Methane oxidation: A process which reduces methane in the environment, thereby reducing warming.

Methyl groups: Small chemical groups made of one carbon atom and three hydrogen atoms ($-CH_3$). They can be added to DNA or to proteins associated with DNA, such as histones.

Methylation: The addition of methyl groups to DNA or proteins. DNA methylation is an epigenetic modification that can help regulate gene expression.

Microbiome: The community of microorganisms (including bacteria, archaea, fungi, protists and viruses) inhabiting a particular environment or living in association with a host organism, together with their collective genetic material.

Microfluidics: The science and technology of manipulating very small volumes of fluids through microminiaturised channels.

Mitochondria: Membrane-bound organelles found in most eukaryotic cells. Their role is to generate chemical energy.

Multiplexing: Measuring several target molecules simultaneously.

Mutualism: A type of relationship between two different species in which both partners benefit from the interaction.

Nanopore sequencing: Detects changes in current as individual nucleic acid molecules pass through a nano-scale pore.

Natural products: Compounds from nature that are useful in food or medicine.

Next Generation Sequencing (NGS): Newer DNA sequencing technologies that sequence millions of small fragments of DNA in parallel. Also referred to as 'massively parallel' sequencing. NGS tends to be cheaper and faster than Sanger sequencing but requires bioinformatic analyses to generate the end sequence.

Nitroplast: A nitrogen-fixing structure within cells.

Non-coding RNA molecules: Those that do not code for proteins and instead have regulatory functions.

Nuclear Magnetic Resonance (NMR): A technique to determine the structure and properties of molecules.

Nucleotide: A molecule that serves as the fundamental building block of nucleic acids such as DNA and RNA.

-Omics: A suite of tools used to characterise and analyse entire sets of molecules (i.e. genes, transcripts, proteins, or metabolites) within an organism or environmental sample.

Operational Taxonomic Unit: A functional proxy for species when using DNA sequence data and is used to group organisms based on their genetic similarity, especially when it is not possible to identify every species precisely.

Organelle: A specialised structure (e.g. the nucleus, mitochondria or chloroplasts) within a cell that carries out a particular function.

Otoliths: Small calcium carbonate structures in the inner ear of bony fish. They grow in layers throughout the life of the fish, and the resulting growth rings are traditionally counted to estimate age.

Oxygen minimum zone: A region of the Ocean that naturally has low levels of dissolved oxygen and is consequently unsuitable habitat for many animals.

Passive samplers: Collect biological material from water over time without the use of mechanical pumps.

Phycosphere: The microenvironment surrounding plankton cells that harbours distinct bacterial communities involved in metabolic exchanges with their algal hosts.

Phylogenetic analyses: These are carried out to determine how closely related different organisms are.

Physiological plasticity: The ability to adapt functions and processes in response to changes in the environment.

Phytoplankton: Microscopic plants or algae that live in the water column and are important primary producers in aquatic ecosystems.

Polymerase Chain Reaction (PCR): A laboratory technique used to amplify copies of specific DNA sequences, generally present in small amounts, thus facilitating their detection and sequence analysis.

Polymetallic nodules: Rocks on the deep-sea floor that contain natural deposits of minerals including manganese and iron.

Primer: A short piece of DNA that adheres to a specific section of DNA, acting as a starting point for DNA synthesis.

Prokaryote: Organisms (i.e. Bacteria and Archaea) whose cells lack a membrane-bound nucleus and membrane-bound organelles.

Proteins: Molecules that play essential roles in the structure, function and regulation of organisms. Encoded by genes and formed by amino acids, these are the building blocks of life.

Proteomics: The study of proteins produced by a single cell, tissue, organism, or ecosystem.

Protist: A diverse group of eukaryotic single-celled (or colonial) organisms, such as unicellular algae, ciliates or amoebas, not classifiable as animals, plants, or fungi.

Quantitative PCR (qPCR): A laboratory technique to amplify and simultaneously quantify a specific DNA or RNA marker in real-time.

RDF triples: A standard format for representing data and relationships on the internet and in semantic databases.

Reference databases: Curated digital libraries of information about biomolecules including their sequences, functions and the names of the organisms they belong to.

Resting cyst: A dormant structure formed by some microorganisms to survive adverse environmental conditions. structure.

Ribonucleic Acid (RNA): Molecules that play central roles in protein synthesis, gene regulation and cellular communication. RNA may also carry genetic code in some viruses.

RNA interference: A process by which small RNA molecules inhibit gene expression.

RNA sequencing (RNAseq): A high-throughput technique used to quantify the complete set of RNA transcripts in a sample to determine which genes are expressed and their (relative) expression levels.

RPA-CRISPR: A laboratory technique that combines Recombinase Polymerase Amplification (RPA) with CRISPR-Cas12a technology to provide a sensitive diagnostic tool that can provide real-time detection of pathogens.

Sanger sequencing: A DNA sequencing method (also known as chain termination sequencing) that uses modified nucleotides to terminate DNA synthesis at specific points.

Secondary metabolites: Chemicals produced by organisms that are not directly involved in essential processes but have functions such as defence, competition or communication.

Semantically structured data: Data that are organised with clear labels that explain exactly what the data mean and how they connect to other data.

Ships of opportunity: Existing vessels, such as commercial ships, ferries or fishing boats, that can be used to collect environmental samples or observations during their normal operations.

Short-read technologies: A technique by which the sample DNA is fragmented into short lengths prior to sequencing and reads of 50-300 base pairs are later bioinformatically assembled into longer contigs.

Single Nucleotide Polymorphism (SNP): A single-letter difference in the DNA sequence among individuals or populations. SNPs are commonly used as genetic markers.

Species: The lowest taxonomic unit of biological classification. A species is a lineage of organisms, based on reproductive compatibility, evolutionary lineage, morphology, ecology, genetics, or a combination of evidence.

Standard curves: Charts made using DNA from reference genes at known concentrations vs fluorescence signal, to calculate the amount of DNA in the test sample.

Standard Operating Procedures: Detailed, step-by-step protocols for sampling, laboratory analysis, and reporting designed to harmonise methods and ensure reproducibility across programs and sites.

Symbiont: An organism that lives in a close, long-term association with another species, the host. Symbionts may benefit the host, as in mutualism, be neutral, as in commensalism, or be harmful, as in parasitism.

Taxon (pl. Taxa): A group of organisms that is classified together in taxonomy. Each taxon represents a unit at any rank in the taxonomic hierarchy, such as species, genus, family, order, class, phylum, or kingdom.

Taxonomy: The naming, describing, and classification of organisms into hierarchical groups based on shared traits and evolutionary relationships.

Transcription: The process by which information encoded in DNA is copied into RNA.

Transcriptomics: The study of transcripts of a cell, tissue or organism across a variety of different biological conditions, which provides information on gene expression.

Trophic: The feeding and nutrition relationships among organisms in an ecosystem.

Ultra Conserved Elements: Segments of DNA that are identical or nearly identical across different species.

Viroid: A viroid is a small, infectious agent made up of a short strand of RNA with no protein shell.

References

Agersnap, et al., (2022). A National Scale "BioBlitz" Using Citizen Science and eDNA Metabarcoding for Monitoring Coastal Marine Fish. *Frontiers in Marine Science, 9*. https://doi.org/10.3389/fmars.2022.824100

Ahmed, W. et al., (2019). Comparative decay of sewage-associated marker genes in beach water and sediment in a subtropical region. *Water Research, 149,* 511–521. https://doi.org/10.1016/j.watres.2018.10.088

Albertsen, M. et al., (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnology, 31*(6), 533–538. https://doi.org/10.1038/nbt.2579

Alrefaey, A. I., et al., (2026). A lab-on-a-chip system integrating DNA purification and loop-mediated isothermal amplification for the quantification of the toxic diatom Pseudo-nitzschia multistriata. *Sensors and Diagnostics, 5*(1), 63–75. https://doi.org/10.1039/d5sd00135h

Anwar, M. Z., et al., (2019). To assemble or not to resemble—A validated Comparative Metatranscriptomics Workflow (CoMW). *GigaScience, 8*(8), giz096. https://doi.org/10.1093/gigascience/giz096

Armbrecht, L. (2020). The Potential of Sedimentary Ancient DNA to Reconstruct Past Ocean Ecosystems. *Oceanography, 33*(2). https://doi.org/10.5670/oceanog.2020.211

Austen, M. C., et al., (2019). Valuing Marine Ecosystems - Taking into account the value of ecosystem benefits in the Blue Economy. In J. Coopman, J. J. Heymans, P. Kellett, A. Muñiz Piniella, V. French, & B. Alexander (Eds.), *EMB Future Science Brief 5.* https://doi.org/10.5281/zenodo.260273

Avsec, Ž., et al., (2026). Advancing regulatory variant effect prediction with AlphaGenome. *Nature, 649*(8099), 1206–1218. https://doi.org/10.1038/s41586-025-10014-0

Aylagas, E., et al., (2020). Translational Molecular Ecology in practice: Linking DNA-based methods to actionable marine environmental management. *Science of The Total Environment, 744,* 140780. https://doi.org/10.1016/j.scitotenv.2020.140780

Bar-On, Y. M., et al., (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences, 115*(25), 6506–6511. https://doi.org/10.1073/pnas.1711842115

Belkin, N., et al., (2026). Introducing Microalgae Carbon Fixation and Sinking (MCFS): a new approach for controlled and scalable CDR. Cambridge Prisms: Carbon Technologies (Preprint). https://eartharxiv.org/repository/view/12519/

Benedetti-Cecchi, L., et al., (2018). Strengthening Europe's Capability in Biological Ocean Observations. In A. Muniz Piniella, P. Kellett, K. Larkin, & J. J. Heymans (Eds.), *EMB Future Science Brief 3.* http://www.marineboard.eu/publication/strengthening-europes-capability-biological-ocean-observations-future-science-brief

Benedicenti, O., et al., (2026). Integrated approaches for pathogen monitoring and shotgun metagenomic analysis in Atlantic salmon farming. *Scientific Reports, 16*(1), 17468. https://doi.org/10.1038/s41598-026-48791-x

Bergauer, K., et al., (2018). Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. *Proceedings of the National Academy of Sciences, 115*(3), E400–E408. https://doi.org/10.1073/pnas.1708779115

Berry, O., et al., (2021). Making environmental DNA (eDNA) biodiversity records globally accessible. *Environmental DNA, 3*(4), 699–705. https://doi.org/10.1002/edn3.173

Bhendarkar, et al., (2025). Advancing ecological assessment: The integration of eDNA metabarcoding into an estuarine fish index. BioRxiv. https://doi.org/10.1101/2025.03.31.645977

Blancher, P., et al., (2022). A strategy for successful integration of DNA-based methods in aquatic monitoring. *Metabarcoding and Metagenomics, 6*, e85652. https://doi.org/10.3897/mbmg.6.85652

Blanco-Bercial, L., et al., (2025). MetaZooGene Intercalibration Experiment (MZG-ICE): Metabarcoding Marine Zooplankton Diversity of the Global Ocean. *Molecular Ecology Resources, 26*(1), e70090. https://doi.org/10.1111/1755-0998.70090

Borja, A., et al., (2024a). Addressing the cumulative impacts of multiple human pressures in marine systems, for the sustainable use of the seas. *Frontiers in Ocean Sustainability, 1.* https://doi.org/10.3389/focsu.2023.1308125

Borja, A., et al., (2024b). DNA-based marine benthic assessment methods can perform as morphological ones, but an intercalibration is needed. *Ecological Indicators, 167,* 112638. https://doi.org/10.1016/j.ecolind.2024.112638

Bosch-Belmar, M., et al., (2021). Jellyfish Impacts on Marine Aquaculture and Fisheries. *Reviews in Fisheries Science & Aquaculture, 29*(2), 242–259. https://doi.org/10.1080/23308249.2020.1806201

Bouchet, P., et al., (2023). Marine biodiversity discovery: the metrics of new species descriptions. F*rontiers in Marine Science, 10*. https://doi.org/10.3389/fmars.2023.929989

Brennan, R. S., et al., (2022). Loss of transcriptional plasticity but sustained adaptive capacity after adaptation to global change conditions in a marine copepod. *Nature Communications, 13*(1), 1147. https://doi.org/10.1038/s41467-022-28742-6

Callens, M., et al., (2025). An Accessible Metagenomic Strategy Allows for Better Characterisation of Invertebrate Bulk Samples. *Molecular Ecology Resources, 25*(8), e14126. https://doi.org/10.1111/1755-0998.14126

Cañas, B., et al., (2007). Trends in sample preparation for classical and second generation proteomics. *Journal of Chromatography A, 1153*(1), 235–258. https://doi.org/10.1016/j.chroma.2007.01.045

CBD (2022). Kunming-Montreal Global Biodiversity Framework: Decision adopted by the Conference of the Parties to the Convention on Biological Diversity (CBD/COP/DEC/15/4). *United Nations Environment Programme.* https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf

CBD (2024). Digital sequence information on genetic resources: Decision adopted by the Conference of the Parties to the Convention on Biological Diversity on 1 November 2024 (CBD/COP/DEC/16/2). *United Nations Environment Programme.* https://www.cbd.int/doc/decisions/cop-16/cop-16-dec-02-en.pdf

Celemín, E., et al., (2025). Evolutionary history and seascape genomics of Harbour porpoises (Phocoena phocoena) across environmental gradients in the North Atlantic and adjacent waters. *Molecular Ecology Resources, 25*(5), e13860. https://doi.org/10.1111/1755-0998.13860

Chaffron, S., et al., (2021). Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Science Advances, 7*(35), eabg1921. https://doi.org/10.1126/sciadv.abg1921

Christodoulou, M., et al., (2025). A taxonomically reliable DNA barcode reference library for North Sea macrobenthos. *Scientific Data, 12*(1), 1198. https://doi.org/10.1038/s41597-025-05500-z

United Nations (2023). Agreement under the United Nations Convention on the Law of the Sea on the Conservation and Sustainable Use of Marine Biological Diversity of Areas beyond National Jurisdiction (A/CONF.232/2023/4). United Nations Digital Library. https://digitallibrary.un.org/record/4013344?ln=en

Coale, T. H., et al., (2024). Nitrogen-fixing organelle in a marine alga. *Science, 384*(6692), 217–222. https://doi.org/10.1126/science.adk1075

COM/2020/380 final. (2020). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: EU Biodiversity Strategy for 2030 Bringing nature back into our lives. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52020DC0380

COM/2025/ 3 final. (2025). Report from the Commission to the Council and the European Parliament on the Commission's assessment of the Member State's programmes of measures as updated under Article 17 of the Marine Strategy Framework Directive (2008/56/EC). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025DC0003

COM/2025/281 final. (2025). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: The European Ocean Pact. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025DC0281

COM/2026/268 final. (2026). Communication from the Commission to the European Parliament and the Council: OceanEye: An EU ocean observation initiative. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:52026DC0268

Commission Implementing Regulation (EU) 2019/627. (2019). Commission Implementing Regulation (EU) 2019/627 of 15 March 2019 laying down uniform practical arrangements for the performance of official controls on products of animal origin intended for human consumption in accordance with Regulation (EU) 2017/625 of the European Parliament and of the Council and amending Commission Regulation (EC) N° 2074/2005 as regards official controls. *Official Journal of the European Union, 131,* 51 - 100. https://eur-lex.europa.eu/eli/reg_impl/2019/627/oj/eng

Conesa, A., et al., (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology, 17*(1), 13. https://doi.org/10.1186/s13059-016-0881-8

Cordier, T., et al., (2021). Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. *Molecular Ecology, 30*(13), 2937–2958. https://doi.org/10.1111/mec.15472

Cordier, T., et al., (2019). Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. *Trends in Microbiology, 27*(5), 387–397. https://doi.org/10.1016/j.tim.2018.10.012

Council Directive 92/43/EEC. (1992). Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. *Official Journal of the European Union, 11,* 0114. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31992L0043

Daraghmeh, N., et al., (2025). A Long-Term Ecological Research Data Set From the Marine Genetic Monitoring Program ARMS-MBON 2018–2020. *Molecular Ecology Resources, 25*(4), e14073. https://doi.org/10.1111/1755-0998.14073

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society, 85*(3), 407–417. https://doi.org/10.1111/j.1095-8312.2005.00503.x

DeLong, E. F. (2009). The microbial ocean from genomes to biomes. *Nature, 459*(7244), 200–206. https://doi.org/10.1038/nature08059

Directive 91/271/EEC. (1991). Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment. *Official Journal of the European Union, 135,* 40–52. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31991L0271

Directive 2000/60/EC. (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Union, 327,* 0001–0037. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0060

Directive 2006/7/EC. (2006). Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. *Official Journal of the European Union, 64,* 37 - 51. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32006L0007

Directive 2008/56/EC. (2008). Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). *Official Journal of the European Union, L164/19.* https://eur-lex.europa.eu/eli/dir/2008/56/oj/eng

Dittami, S., et al., (2021). A community perspective on the concept of marine holobionts: current status, challenges, and future directions. PeerJ, 9. https://doi.org/10.7717/peerj.10911

Drouet, K., et al., (2021). Current distribution and potential expansion of the harmful benthic dinoflagellate Ostreopsis cf. siamensis towards the warming waters of the Bay of Biscay, North-East Atlantic. *Environmental Microbiology, 23*(9), 4956–4979. https://doi.org/10.1111/1462-2920.15406

El Hourany, R., et al., (2024). Linking satellites to genes with machine learning to estimate phytoplankton community structure from space. *Ocean Science, 20*(1), 217–239. https://doi.org/10.5194/os-20-217-2024

Environmental Genomics Research Programme. (2025). Efficacy of eDNA vs Conventional Monitoring Methods. https://www.iogp-edna.org/wp-content/uploads/2026/02/Chapter-1-Efficacy-of-eDNA-vs-Conventional-Monitoring-Methods-June-2025.pdf

Esteban-Cantillo, O. J., et al., (2024). Dynamic Oceans, Dynamic Solutions. BiOCean5D GA# 101059915. 2024. hal-04867560. https://hal.science/hal-04867560v1

European Marine Board. (2019). Navigating the Future V: Marine Science for a Sustainable Future. In J. J. Heymans, B. Alexander, A. Muniz Piniella, P. Kellett, J. Coopman, & K. Larkin (Eds.), *EMB Position Paper 24.* https://doi.org/10.5281/zenodo.2809392

Falkowski, P. G., et al., (2008). The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science, 320*(5879), 1034–1039. https://doi.org/10.1126/science.1153213

Fraija-Fernández, N., et al., (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution, 10*(14), 7560–7584. https://doi.org/10.1002/ece3.6482

Fukuba, T., & Fujii, T. (2006). Development and Evaluation of the Integrated *In Situ* Analyzer for Gene-"IISA-Gene" for Microbiology in Extreme Environments. 2006 International Conference on Microtechnologies in Medicine and Biology, 184–187. https://doi.org/10.1109/MMB.2006.251523

Garcia Soto, C., et al., (2017). Advancing Citizen Science for Coastal and Ocean Research. In V. French, P. Kellett, J. Delany, & N. McDonough (Eds.), *EMB Position Paper 23*. http://marineboard.eu/publication/advancing-citizen-science-coastal-and-ocean-research-position-paper

Giordano, N., et al., (2024). Genome-scale community modelling reveals conserved metabolic cross-feedings in epipelagic bacterioplankton communities. *Nature Communications, 15*(1), 2721. https://doi.org/10.1038/s41467-024-46374-w

Gomez Cortes, L., et al., (2024). The European trial exercise on Escherichia coli detection in bathing waters. (KJ-NA-31-969-EN-N (online). https://doi.org/10.2760/42220

Grillo, M., et al., (2024). Machine learning applied to species occurrence and interactions: the missing link in biodiversity assessment and modelling of Antarctic plankton distribution. *Ecological Processes, 13*(1), 56. https://doi.org/10.1186/s13717-024-00532-6

Guérin, N., et al., (2022). Genomic adaptation of the picoeukaryote Pelagomonas calceolata to iron-poor oceans revealed by a chromosome-scale genome sequence. *Communications Biology, 5*(1), 983. https://doi.org/10.1038/s42003-022-03939-z

Guidi, L., et al., (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature, 532*(7600), 465–470. https://doi.org/10.1038/nature16942

Guidi, L., et al., (2020). Big Data in Marine Science. In J. J. Heymans, B. Alexander, A. Muniz Piniella, P. Kellett, & J. Coopman (Eds.), *EMB Future Science Brief 6.* https://doi.org/10.5281/zenodo.3755793

Harðardóttir, S., et al., (2024). Millennial-scale variations in Arctic sea ice are recorded in sedimentary ancient DNA of the microalga Polarella glacialis. *Communications Earth & Environment, 5*(1), 25. https://doi.org/10.1038/s43247-023-01179-5

He, C., et al., (2021). Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nature Microbiology, 6*(3), 354–365. https://doi.org/10.1038/s41564-020-00840-5

Humphries, F. (2025). Marine Genetic Resources Beyond National Jurisdiction: The Expansive Scope of the BBNJ Agreement. In F. Humphries (Ed.), Decoding Marine Genetic Resource Governance Under the BBNJ Agreement (pp. 55–94). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72100-7_3

Hutchins, D. A., & Fu, F. (2017). Microorganisms and ocean global change. *Nature Microbiology, 2*(6), 17058. https://doi.org/10.1038/nmicrobiol.2017.58

Inoue, J., et al., (2026). eDNAmap: A Metabarcoding Web Tool for Comparing Marine Biodiversity, With Special Reference to Teleost Fish. *Molecular Ecology Resources, 26*(1), e70066. https://doi.org/10.1111/1755-0998.70066

IPBES. (2019). Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES Secretariat. https://doi.org/10.5281/ZENODO.6417333

Jo, T., & Minamoto, T. (2021). Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta-analyses. *Molecular Ecology Resources, 21*(5), 1490–1503. https://doi.org/10.1111/1755-0998.13354

John, U., et al., (2026). Towards standardized molecular monitoring of HABs: A consensus guide to qPCR. *Harmful Algae, 158,* 103148. https://doi.org/10.1016/j.hal.2026.103148

Kaneko, H., et al., (2023). Predicting global distributions of eukaryotic plankton communities from satellite data. *ISME Communications, 3*(1), 101. https://doi.org/10.1038/s43705-023-00308-7

Kavakiotis, I., et al., (2026). Navigating the European aquatic eDNA landscape: Opportunities for metadata standardisation and data mobilisation. *Metabarcoding and Metagenomics, 10,* e173612. https://doi.org/10.3897/mbmg.10.173612

Kim, K., et al., (2024). Detection of the jellyfish Chrysaora pacifica by RPA-CRISPR-Cas12a environmental DNA (eDNA) assay and its evaluation through field validation and comparative eDNA analyses. *Science of The Total Environment, 955,* 176945. https://doi.org/10.1016/j.scitotenv.2024.176945

Koonin, E. V., & Lee, B. D. (2025). Diversity and evolution of viroids and viroid-like agents with circular RNA genomes revealed by metatranscriptome mining. *Nucleic Acids Research, 53*(3), gkae1278. https://doi.org/10.1093/nar/gkae1278

Kornienko, I. V, et al., (2024). RNA Stability: A Review of the Role of Structural Features and Environmental Conditions. *Molecules, 29*(24). https://doi.org/10.3390/molecules29245978

Kujawinski, E. B. (2011). The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annual Review of Marine Science, 3*, 567–599. /https://doi.org/10.1146/annurev-marine-120308-081003

Lai, K. P., et al., (2016). Hypoxia alters steroidogenesis in female marine medaka through miRNAs regulation. Aquatic Toxicology, 172, 1–8. https://doi.org/10.1016/j.aquatox.2015.12.012

Lambert, A., et al., (2024). Community metabolic modeling of host-microbiota interactions through multi-objective optimization. *IScience, 27*(6). https://doi.org/10.1016/j.isci.2024.110092

Laroche, O., et al., (2020). Environmental DNA surveys detect distinct metazoan communities across abyssal plains and seamounts in the western Clarion Clipperton Zone. *Molecular Ecology, 29*(23), 4588–4604. https://doi.org/10.1111/mec.15484

Laudadio, I., et al., (2019). Next-Generation Metagenomics: Methodological Challenges and Opportunities. *OMICS: A Journal of Integrative Biology, 23*(7), 327–333. https://doi.org/10.1089/omi.2019.0073

Lee, H.-G., et al., (2020). Quantification of the paralytic shellfish poisoning dinoflagellate Alexandrium species using a digital PCR. *Harmful Algae, 92,* 101726. https://doi.org/10.1016/j.hal.2019.101726

Lee, Y., et al., (2022). Oxidative stress–mediated synergistic deleterious effects of nano- and microplastics in the hypoxia-conditioned marine rotifer Brachionus plicatilis. *Marine Pollution Bulletin, 181,* 113933. https://doi.org/10.1016/j.marpolbul.2022.113933

Leese, F., et al., (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes, 2,* e11321. https://doi.org/10.3897/rio.2.e11321

Lehmann, I., et al., (2026). Implications of global distributive justice principles for implementation of the Kunming-Montreal Global Biodiversity Framework. *Conservation Biology, 40*(1), e70167. https://doi.org/10.1111/cobi.70167

Lombard, F., et al., (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Frontiers in Marine Science, Volume 6-2019*. https://doi.org/10.3389/fmars.2019.00196

Lumbierres, M., et al., (2024). EuropaBON EBV workflow templates. Zenodo. https://doi.org/10.5281/zenodo.10971094

Maiello, G., et al., (2022). Little samplers, big fleet: eDNA metabarcoding from commercial trawlers enhances ocean monitoring. *Fisheries Research, 249,* 106259. https://doi.org/10.1016/j.fishres.2022.106259

Martiny, A. C., et al., (2026). Genomic-to-space measurements reveal large-scale ocean nutrient stress. *Science Advances, 12*(23), eaed8089. https://doi.org/10.1126/sciadv.aed8089

Mason, O. U., et al., (2012). Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *The ISME Journal, 6*(9), 1715–1727. https://doi.org/10.1038/ismej.2012.59

Metfies, K., et al., (2005). Electrochemical detection of the toxic dinoflagellate Alexandrium ostenfeldii with a DNA-biosensor. *Biosensors and Bioelectronics, 20*(7), 1349–1357. https://doi.org/10.1016/j.bios.2004.05.011

Mo, Y., et al., (2026). A Beginner's Guide to Using DeepVirFinder for Viral Sequence Identification From Metagenomic Datasets. *Current Protocols, 6*(2), e70310. https://doi.org/10.1002/cpz1.70310

Moltmann, T., et al., (2019). A Global Ocean Observing System (GOOS), Delivered Through Enhanced Collaboration Across Regions, Communities, and New Technologies. *Frontiers in Marine Science, 6.* https://doi.org/10.3389/fmars.2019.00291

Moran, M. A., et al., (2022). Microbial metabolites in the marine carbon cycle. *Nature Microbiology, 7(*4), 508–523. https://doi.org/10.1038/s41564-022-01090-3

Moran, M. A., et al., (2016). Deciphering ocean carbon in a changing world. *Proceedings of the National Academy of Sciences, 113*(12), 3143–3151. https://doi.org/10.1073/pnas.1514645113

Moran, M. A., et al., (2013). Sizing up metatranscriptomics. *The ISME Journal, 7*(2), 237–243. https://doi.org/10.1038/ismej.2012.94

Mugo, R., & Saitoh, S.-I. (2020). Ensemble Modelling of Skipjack Tuna (Katsuwonus pelamis) Habitats in the Western North Pacific Using Satellite Remotely Sensed Data; a Comparative Analysis Using Machine-Learning Models. *Remote Sensing, 12*(16). https://doi.org/10.3390/rs12162591

Muhling, B. A., et al., (2020). Predictability of Species Distributions Deteriorates Under Novel Environmental Conditions in the California Current System. *Frontiers in Marine Science, Volume 7-2020.* https://doi.org/10.3389/fmars.2020.00589

Muri, H., et al., (2025). Monitoring, reporting and verification for marine carbon dioxide removal. In: A. Muñiz Piniella, A. Rodriguez Perez, P. Kellett, B. Alexander, F. Bayo Ruiz, & J. J. Heymans, Eds. European Marine Board. https://doi.org/10.5281/zenodo.17435116

Nikolaou, A., et al., (2025). What Do We Know About the Environmental Status of European Seas? *Conservation Letters, 18.* https://doi.org/10.1111/conl.13118

Oliver, D. M., et al., (2016). Molecular tools for bathing water assessment in Europe: Balancing social science research with a rapidly developing environmental science evidence-base. *Ambio, 45*(1), 52–62. https://doi.org/10.1007/s13280-015-0698-9

Pagnier, J., et al., (2025). Using the long-term genetic monitoring network ARMS-MBON to detect marine non-indigenous species along the European coasts. *Biological Invasions, 27*(2), 77. https://doi.org/10.1007/s10530-024-03503-2

Palialexis, A., et al., (2021). Monitoring biodiversity for the EU Marine Strategy Framework Directive: Lessons learnt from evaluating the official reports. *Marine Policy, 128,* 104473. https://doi.org/10.1016/j.marpol.2021.104473

Paoli, L., et al., (2022). Biosynthetic potential of the global ocean microbiome. *Nature, 607*(7917), 111–118. https://doi.org/10.1038/s41586-022-04862-3

Patti, G. J., et al., (2012). Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology, 13*(4), 263–269. https://doi.org/10.1038/nrm3314

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

Petit-Marty, N., et al., (2025). Genomic patterns and levels of genetic diversity reveal recent declines in the effective size of European hake populations. *ICES Journal of Marine Science, 82*(12), fsaf231. https://doi.org/10.1093/icesjms/fsaf231

Pinfield, R., et al., (2019). False-negative detections from environmental DNA collected in the presence of large numbers of killer whales (Orcinus orca). *Environmental DNA, 1*(4), 316–328. https://doi.org/10.1002/edn3.32

Rabone, M., et al., (2023). How many metazoan species live in the world's largest mineral exploration region? *Current Biology, 33*(12), 2383-2396.e5. https://doi.org/10.1016/j.cub.2023.04.052

Raes, E., et al., (2026). Catalysing the uptake of environmental DNA (eDNA) into marine park management across Australia. *Zenodo.* https://doi.org/10.5281/zenodo.21929550

Régimbeau, A., et al., (2025). Unveiling the link between phytoplankton molecular physiology and biogeochemical cycling via genome-scale modeling. *Science Advances, 1*1(11), eadq3593. https://doi.org/10.1126/sciadv.adq3593

Regulation EU 1380/2013. (2013). Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy (2013). *Official Journal of the European Union, 354, 22.* https://eur-lex.europa.eu/eli/reg/2013/1380/oj/eng

Regulation (EU) 2024/1991. (2024). Regulation (EU) 2024/1991 of the European Parliament and of the Council of 24 June 2024 on nature restoration and amending Regulation (EU) 2022/869. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1991&qid=1722240349976

Ribeiro, H., et al., (2019). Development of an autonomous biosampler to capture *in situ* aquatic microbiomes. *PLOS ONE, 14*(5), e0216882-. https://doi.org/10.1371/journal.pone.0216882

Richter, D. J., et al., (2022). Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. *ELife, 11,* e78129. https://doi.org/10.7554/eLife.78129

Rimet, F., et al., (2026). eDNAqua-Plan Deliverable 9: Recommendations for a federated system of taxonomically curated reference libraries. *Zenodo.* https://doi.org/10.5281/zenodo.18849393

Ríos-Castro, R., et al., (2023). Harmful algae diversity from a coastal upwelling system detected by high-throughput sequencing. *Frontiers in Marine Science, 10.* https://doi.org/10.3389/fmars.2023.1200135

Rodrigues, P. M., et al., (2018). Proteomics in Fish and Aquaculture Research. In A. M. de Almeida, D. Eckersall, & I. Miller (Eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology* (pp. 311–338). Springer International Publishing. https://doi.org/10.1007/978-3-319-69682-9_16

Rourke, M. L., et al., (2022). Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA, 4*(1), 9–33. https://doi.org/10.1002/edn3.185

Ruvindy, R., et al., (2018). qPCR Assays for the Detection and Quantification of Multiple Paralytic Shellfish Toxin-Producing Species of Alexandrium. *Frontiers in Microbiology, Volume 9-2018.* https://doi.org/10.3389/fmicb.2018.03153

Salter, I., et al., (2019). Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. *Communications Biology, 2*(1), 461. https://doi.org/10.1038/s42003-019-0696-8

Sander, S., et al., (2025). Deep Sea Research and Management Needs (p. 80). *Zenodo.* https://doi.org/10.5281/zenodo.14928917

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

Santi, I., et al., (2021). European Marine Omics Biodiversity Observation Network (EMO BON) Handbook (Version 1.0). EMBRC-ERIC. http://dx.doi.org/10.25607/OBP-1653

Santi, I., et al., (2021). A Comparison of DNA Metabarcoding and Microscopy Methodologies for the Study of Aquatic Microbial Eukaryotes. *Diversity, 13*(5). https://doi.org/10.3390/d13050180

Sarmento, H., et al., (2025). The southern gap in ocean microbiome science. *Ocean Microbiology, 1*(1), 6. https://doi.org/10.1186/s44375-025-00006-w

Schickele, A., et al., (2025). CEPHALOPOD, a package to standardize marine habitat-modelling practices and enhance inter-comparability across biological observations. *Methods in Ecology and Evolution, 16*(6), 1126–1135. https://doi.org/10.1111/2041-210X.70040

Scholin, C., et al., (2017). The Quest to Develop Ecogenomic Sensors: A 25-Year History of the Environmental Sample Processor (ESP) as a Case Study. *Oceanography, 30,* 100–113. https://doi.org/10.5670/oceanog.2017.427

Serandour, B., et al., (2026). Satellite-gene abundance fusion reveals global hotspots of ocean ecosystem services. BioRxiv. https://doi.org/10.64898/2026.09.23.753815

Shelton, A. O., et al., (2022). Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. *Proceedings of the Royal Society B: Biological Sciences, 289*(1971), 20212613. https://doi.org/10.1098/rspb.2021.2613

Sher, D., et al., (2024). Quantitative principles of microbial metabolism shared across scales. *Nature Microbiology, 9*(8), 1940–1953. https://doi.org/10.1038/s41564-024-01764-0

Shilova, I. N., et al., (2014). A microarray for assessing transcription from pelagic marine microbial taxa. *The ISME Journal, 8*(7), 1476–1491. https://doi.org/10.1038/ismej.2014.1

Siano, R., et al., (2021). Sediment archives reveal irreversible shifts in plankton communities after World War II and agricultural pollution. *Current Biology, 31*(12), 2682-2689.e7. https://doi.org/10.1016/j.cub.2021.03.079

Sunagawa, S., et al., (2020). Tara Oceans: towards global ocean ecosystems biology. *Nature Reviews Microbiology, 18*(8), 428–445. https://doi.org/10.1038/s41579-020-0364-5

Sunagawa, S., et al., (2015). Structure and function of the global ocean microbiome. *Science, 348*(6237), 1261359. https://doi.org/10.1126/science.1261359

Sushmitha, T. J., et al., (2025). Exploring the abyss through recent advances in microbial profiling technologies. *Discover Oceans, 2*(1), 68. https://doi.org/10.1007/s44289-025-00075-4

Tagliabue, A. (2023). 'Oceans are hugely complex': modelling marine microbes is key to climate forecasts. *Nature, 623*(7986), 250–252. https://doi.org/10.1038/d41586-023-03425-4

Thangaraj, S., & Sun, J. (2026). Omics insights into ocean health: Molecular adaptations and ecosystem resilience under climate stress. *Advances in Climate Change Research, 17*(1), 199–217. https://doi.org/10.1016/j.accre.2025.10.004

Tomanek, L. (2011). Environmental Proteomics: Changes in the Proteome of Marine Organisms in Response to Environmental Stress, Pollutants, Infection, Symbiosis, and Development. *Annual Review of Marine Science, 3,* 373–399. https://doi.org/10.1146/annurev-marine-120709-142729

Underwood, J. N., et al., (2018). Genetic signatures through space, time and multiple disturbances in a ubiquitous brooding coral. *Molecular Ecology, 27*(7), 1586–1602. https://doi.org/10.1111/mec.14559

United Nations. (2026). The Third World Ocean Assessment (WOA III). Division for Ocean Affairs and the Law of the Sea, Office of Legal Affairs. https://woa.un.org

Ussler, W., et al., (2024). Underway measurement of cyanobacterial microcystins using a surface plasmon resonance sensor on an autonomous underwater vehicle. *Limnology and Oceanography: Methods, 22*(9), 681–699. https://doi.org/10.1002/lom3.10627

Varadi, M., et al., (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research, 50*(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061

Varrella, S., et al., (2025). A comprehensive assessment of non-indigenous species requires the combination of multi-marker eDNA metabarcoding with classical taxonomic identification. *Environment International, 199,* 109489. https://doi.org/10.1016/j.envint.2025.109489

Wang, F., et al., (2022). CFM-ID 4.0 – a web server for accurate MS-based metabolite identification. *Nucleic Acids Research, 50*(W1), W165–W174. https://doi.org/10.1093/nar/gkac383

Wang, Z., et al., (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics, 10*(1), 57–63. https://doi.org/10.1038/nrg2484

Weber, D. N., et al., (2024). Epigenetic age estimation in a long-lived, deepwater scorpionfish: insights into epigenetic clock development. *Canadian Journal of Fisheries and Aquatic Sciences, 81*(5), 620–631. https://doi.org/10.1139/cjfas-2023-0296

Wegley, K. L., et al., (2018). Ecosystem Microbiology of Coral Reefs: Linking Genomic, Metabolomic, and Biogeochemical Dynamics from Animal Symbioses to Reefscape Processes. *MSystems, 3*(2), 10.1128/msystems.00162-17. https://doi.org/10.1128/msystems.00162-17

WHO. (2021). Guidelines on recreational water quality: Volume 1 Coastal and fresh waters. *World Health Organization, Ed.* https://www.who.int/publications/i/item/9789240031302

Williams, A., et al., (2023). Peeling back the layers of coral holobiont multi-omics data. *IScience, 26*(9), 107623. https://doi.org/10.1016/j.isci.2023.107623

Wollschläger, J., et al., (2014). Assessment of North Sea phytoplankton via molecular sensing: a method evaluation. *Journal of Plankton Research, 36*(3), 695–708. https://doi.org/10.1093/plankt/fbu003

Yamahara, K. M., et al., (2025). A State-Of-The-Art Review of Aquatic eDNA Sampling Technologies and Instrumentation: Advancements, Challenges, and Future Prospects. *Environmental DNA, 7*(4), e70170. https://doi.org/https://doi.org/10.1002/edn3.70170

Yamahara, K. M., et al., (2015). Simultaneous monitoring of faecal indicators and harmful algae using an in-situ autonomous sensor. *Letters in Applied Microbiology, 61*(2), 130–138. https://doi.org/10.1111/lam.12432

Yamahara, K. M., et al., (2019). *In situ* Autonomous Acquisition and Preservation of Marine Environmental DNA Using an Autonomous Underwater Vehicle. *Frontiers in Marine Science, 6,* 373. https://10.3389/fmars.2019.00373

Yang, N., et al., (2024). Applying environmental DNA approaches to inform marine biodiversity conservation: The case of the Ocean Twilight Zone. *Marine Policy, 165,* 106151. https://doi.org/10.1016/j.marpol.2024.106151

Yates, J. R., et al., (2009). Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering, 11*, 49–79. https://doi.org/10.1146/annurev-bioeng-061008-124934

Yooseph, S., et al., (2007). The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families. *PLOS Biology, 5*(3), e16-. https://doi.org/10.1371/journal.pbio.0050016

Zaiko, A., et al., (2021). Towards reproducible metabarcoding data: Lessons from an international cross-laboratory experiment. *Molecular Ecology Resources, 22*(2), 519–538. https://doi.org/10.1111/1755-0998.13485

Zhang, Y., et al., (2019). Targeted Sampling by Autonomous Underwater Vehicles. *Frontiers in Marine Science, 6.* https://doi.org/10.3389/fmars.2019.00415

Zhao, P., et al., (2024). Development of an RPA-CRISPR/Cas12a Assay for Rapid and Sensitive Diagnosis of Plant Quarantine Fungus Setophoma terrestris. *Journal of Fungi, 10*(10). https://doi.org/10.3390/jof10100716

Zimmermann, H. H., et al., (2023). Marine ecosystem shifts with deglacial sea-ice loss inferred from ancient DNA shotgun sequencing. *Nature Communications, 14*(1), 1650. https://doi.org/10.1038/s41467-023-36845-x

Annex 1

Members of the European Marine Board Working Group

NAME	AFFILIATION	COUNTRY
Working Group Chairs		
Louise Allcock	University of Galway	Ireland
Paloma Morán	University of Vigo	Spain
Working Group Members		
Ana Baričević	Ruđer Bošković Institute (IRB)	Croatia
Chris Bowler	Institut de Biologie de l'École Normale Supérieure	France
Damien Eveillard	University of Nantes	France
Pascal Hablützel	Flanders Marine Institute (VLIZ)	Belgium
Bruno Louro	Centro de Ciencias de Mar do Algarve (CCMAR)	Portugal
Antonella Penna	National Interuniversity Consortium for Marine Sciences (CONISMA)	Italy
Benjamin Pontiller	GEOMAR Helmholtz Centre for Ocean Research Kiel	Germany
Julie Robidart	National Oceanography Centre (NOC)	UK
Jorge Juan Santos Bruña	Spanish Institute of Oceanography (IEO-CSIC)	Spain
Elena Stoica	National Institute for Marine Research and Development (NIMRD)	Romania
Agnes Weiner	NORCE Research AS	Norway

Annex 2

AI Tools Used

AI TOOL USED	PURPOSE	CHAPTER/ BOX
Figure Labs	Re-design of Figure 2.2.	Chapter 2
ChatGPT	Re-design image	Box 3.1.
Copilot	Flow of text	Chapter 6
Grammarly	Flow of text	Full document

European
MARINE BOARD
Advancing Seas & Ocean Science

European Marine Board IVZW
Belgian Enterprise Number: 0650.608.890
Jacobsenstraat 1 I 8400 Ostend I Belgium
Tel: +32 (0) 59 33 69 24
E-mail: info@marineboard.eu
www.marineboard.eu