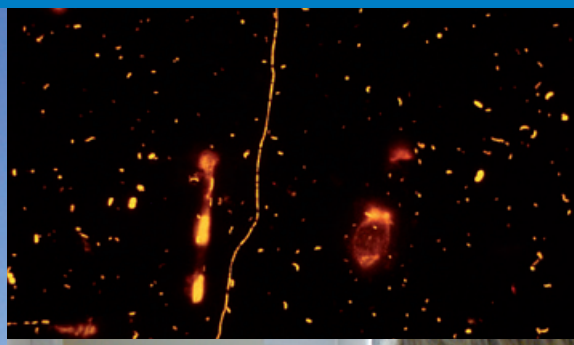
EUROPEAN SCIENCE FOUNDATION

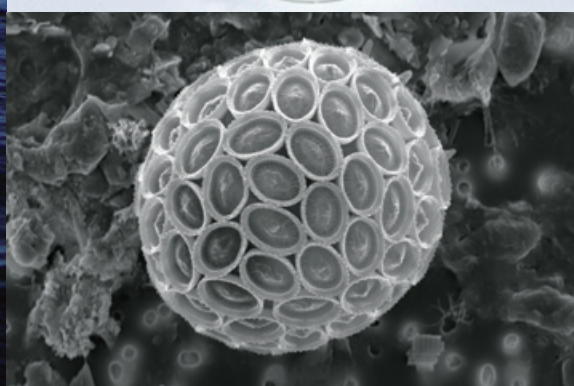
MARINE BOARD

Position Paper 17

Marine Microbial Diversity and its role in Ecosystem Functioning and Environmental Change

May 2012





Marine Board

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

The Marine Board was established in 1995 to facilitate enhanced cooperation between European marine science organisations (both research institutes and research funding agencies) towards the development of a common vision on the research priorities and strategies for marine science in Europe. In 2012, the Marine Board represents 34 Member Organisations from 20 countries. The marine Board provides the essential components for transferring knowledge for leadership in marine research in Europe. Adopting a strategic role, the Marine Board serves its member organisations by providing a forum within which marine research policy advice to national agencies and to the European Commission is developed, with the objective of promoting the establishment of the European Marine Research Area.

www.marineboard.eu

Cover photograph credits:
Left: Research Vessel lowering a rosette with Niskin bottles to obtain seawater samples (courtesy, Lucas Stal)
Right from top to bottom: Fluorescence *in situ* hybridisation picture of a marine water sample stained with a probe specific for Bacteria (courtesy, Frank Oliver Glöckner); Marine microbiologist working in the laboratory (courtesy, Frank Oliver Glöckner and Anna Klindworth); Example of a microorganism culture on a laboratory Petri dish (courtesy, Frank Oliver Glöckner); Image of an alga using Scanning Electron Microscopy (SEM). (courtesy, Ruth-Anne Sandaa).

Marine Microbial Diversity and its role in Ecosystem Functioning and Environmental Change

Marine Board-ESF Position Paper 17

This Position Paper is based on the activities of the Marine Board Working Group on Marine Microbial Diversity (WG MICROCEAN) which convened in Brussels on 17 February 2010, 2-3 September 2010 and 10-11 February 2011, and in Oostende (Belgium) on 15-16 June 2011.

Coordinating author
Frank Oliver Glöckner

Contributing authors
Frank Oliver Glöckner (Chair),
Lucas J. Stal (Coordinator Chapter 2),
Ruth-Anne Sandaa (Coordinator Chapter 3),
Josep Maria Gasol (Coordinator Chapter 4),
Fergal O'Gara, Francisco Hernandez,
Matthias M. Labrenz, Elena Stoica,
Marta Varela Rozados, Adriano Bordalo,
Paraskevi Pitta

Special contributions from
Frede Thingstad, Marcel Kuypers

Series Editor
Niall McDonough

Publication Editor
Jan-Bart Calewaert

Suggested reference:
Glöckner F.O., Stal L.J., Sandaa R.-A., Gasol J.M., O'Gara F., Hernandez F., Labrenz M., Stoica E., Varela M.M., Bordalo A., Pitta P. (2012). Marine Microbial Diversity and its role in Ecosystem Functioning and Environmental Change. Marine Board Position Paper 17. Calewaert, J.B. and McDonough N. (Eds.).
Marine Board-ESF, Ostend, Belgium.

Contents

Foreword

Several Marine Board foresight activities have previously highlighted the importance of marine microbial diversity. In 2007, the Marine Board published its Position Paper 9 on the *Impacts of Climate Change on the European Marine and Coastal Environment*. This publication drew attention to the importance of marine microorganisms for the production of oxygen (approximately 50% of the total photosynthesis on Earth), the removal of carbon dioxide from the atmosphere and the regulation of our climate. The Marine Board Special Report on *Climate Change and Marine Ecosystem Research* (*CLAMER*) published in 2011, also emphasized the key role of microbes in all global cycles of matter and energy and how climate-induced changes in marine microbial communities and interactions may affect critical biogeochemical cycles of elements such as carbon, nitrogen and iron. Meanwhile, the enormous – but largely untapped – potential of marine microorganisms for the development of new biotechnology applications and products was described in Marine Board position paper 15, *Marine Biotechnology: A new Vision and Strategy for Europe* (2010).

In recent years, there has been an increasing recognition of the importance of microbes, coupled with the discovery of a vast microbial diversity. The key role that microbes play in regulating Earth's climate has also emerged as a result of research in this area. However, it is also likely that very rapid (geologically speaking) and ongoing human-induced climate change will, and probably already has, altered the microbial diversity itself. As a result of increasing sea water temperatures, acidification of the ocean and salinity changes, to name just a few environmental parameters of importance, dominant *Bacteria*, *Archaea* and viruses may become dormant and completely unknown species may become dominant. Since we cannot currently predict these changes, it has proven very difficult to predict the microbial influence on biogeochemical cycles and hence factor this influence into predictions of climate change itself.

A number of major European and international projects have made significant progress in addressing marine microbial biodiversity in recent years. With the completion of the FP6 Networks of Excellence, Marine Genomics Europe (MGE) and MarBEF (Marine Biodiversity and Ecosystem Functioning), and large international programmes such as the Census of Marine Life (CoML) and the International Census of Marine Microbes (ICoMM), it is timely to reflect on where we are in the field of marine microbial research and to make plans for the future of European research effort in this area.

Realizing the societal importance of marine microbial diversity and the need to advance and better coordinate the European research effort in this area, the Marine Board set-up a working group of experts (WG MICROCEAN) in 2010 to (i) re-emphasize the importance and role of microbes in the marine environment; (ii) indentify strategic areas for Europe in the realm of marine microbial diversity research with specific reference to ecosystem functioning, biogeochemical cycling and environmental change; (iii) provide recommendations to guide European research in the medium-term (to 2020) to substantially increase our knowledge of marine microbial diversity and its role in, and response to, global change and to improve the competitiveness of European research in this field.

The working group was tasked to deliver a strategic position paper with recommendations designed to directly influence research agendas at both national and European level. In the first instance, this paper is targeted towards those who determine and set the research agenda, including European and national research funding organizations, programme managers and science policy advisors/developers. At the same time, the paper is intended to strengthen the marine microbial research domain by stimulating scientific networking and developing common views of expert scientists, potentially leading to new collaborative projects.

On behalf of the Marine Board, we sincerely thank all of the members of the working group who generously gave their time and expertise to support the production of this important position paper. Their work has been crucial to highlight the importance and central role of microbes in the marine environment and in providing a clear set of research priorities and recommendations to further improve marine microbial research in Europe. Our special thanks goes to the working group Chair, Frank Oliver Glöckner, of the Max Planck Institute for Marine Microbiology, Germany, and to Jan-Bart Calewaert of the Marine Board Secretariat for his diligent support to the working group and his work in finalizing this report. We also thank our external reviewers for their comments and invaluable suggestions, and Ivo Kostadinov for ideas and comments on some parts of the text.

Kostas Nittis,
Chair, Marine Board

Niall McDonough,
Executive Scientific Secretary, Marine Board

Executive summary

Seas and oceans cover more than 70% of the Earth's surface, host the majority of its biomass, and contribute significantly to all global cycles of matter and energy. All life on Earth most likely originated from microbes in the sea. In today's marine ecosystems, following billions of years of evolution, microbes such as *Bacteria*, *Archaea*, viruses, fungi and protists (including microalgae), dominate the living biomass. Recent rapid developments in molecular ecology, metagenomics and ecological modelling illustrate that microbes represent the most important biological group on Earth in terms of phylogenetic and functional diversity. In addition, interdisciplinary research has uncovered new and unexpected roles of microbes in the biogeochemical cycling of carbon, nitrogen, silica and iron and many other (trace) elements in our seas and oceans. Marine microorganisms produce the organic matter and oxygen required to sustain life and facilitate the storage, transport, and turnover of key biological elements. Thus, microorganisms are the foundation of life and are of critical importance to the habitability and sustainability of our planet (Executive Summary Box A).

The enormous microbial diversity also gives rise to a largely untapped amount of genetic information, bioactive compounds and biomaterials which could deliver important benefits and applications of societal interest, for example, to improve medical treatments, fisheries and aquaculture applications, the supply of energy and for the development of industrial products and processes. Yet, despite the clear importance of marine microbes and the major opportunities they present, very little is known about marine microbial diversity, the enormous array of microbial types and their ecological functions and interactions. Moreover, the vast majority (90-99%) of marine microorganisms cannot be cultured under standard laboratory conditions and their growth and physiology cannot, therefore, be studied in the way that has proven so successful throughout the 20th century for medically-important microorganisms.

Addressing these knowledge-gaps will require a significant increase in research investment in Europe, coupled with a better coordination of European researchers, projects, programmes and infrastructures. However, since marine "bugs" fail to capture the public imagination in the same way that, for example, whales or turtles do, making a convincing political case for increased support for marine microbial research can be difficult. Not surprisingly, the non-scientific public is generally unaware of the importance of microbial communities in the functioning and health of our environment. Instead, microbes are most commonly associated with disease in humans, crops and livestock. It is time to educate the public and research policy makers about the importance of marine microbes and marine microbiological research.

Executive Summary Box A.
Key facts highlighting the importance of marine microorganisms

- Marine microorganisms occur in vast numbers and represent a huge genetic diversity: ocean water contains up to one million microorganisms per milliliter and several thousand microbial types. These numbers may be an order of magnitude higher in coastal waters with their higher productivity and higher load of organic matter and nutrients;
- On the tree of life, the *Eukarya* (including plants, animals and protists) comprise only a tiny branch; *Bacteria* and *Archaea* encompass virtually all genetic diversity, notwithstanding their limited morphological diversity;
- Marine microorganisms are key to all biochemical cycles and are crucial for the functioning of marine ecosystems;
- Microbes are responsible for the degradation of organic matter in the ocean and are thus key for maintaining the balance between produced and fixed carbon dioxide;
- Marine phototrophic microorganisms (*Cyanobacteri*a, diatoms and pico- and nanophytoplankton) are responsible for more than 50% of the oxygen production on Earth.
- Marine microorganisms represent a largely untapped source of novel bioactive compounds and metabolic pathways which could be exploited for new biotechnological applications and products;
- Marine microorganisms play an indispensable role in ensuring a sustainable supply of seafood products by occupying the critical bottom trophic level in marine foodwebs and offering solutions for regulating aquaculture processes and bioremediation and waste management;
- Marine microorganisms and their activities are, and will continue to be, affected by global change and may also promote or alleviate climate change.

Europe has traditionally played a leading role in marine microbial research. But the field is advancing rapidly, driven by technological and scientific developments and the problems of global change, population growth and over exploitation of marine resources. With the advent of single cell genomics and meta-omics technologies, we are improving our capacity to assess the status of, and changes in, marine microbial diversity (see also Executive Summary Box B). The enormous expansion in the use of these

techniques has already resulted in a significant increase in the quantity of heterogeneous data which needs to be integrated, analysed and interpreted. At present, Europe has a much smaller capacity than the US in marine data processing and bioinformatics. Furthermore, the existing capacities are often not designed for interoperability within the European research community. It is of crucial importance, therefore, to urgently consider and put in place the bioinformatics infrastructure and standards needed to enable the next generation of European marine microbial research.

Although unprecedented technological improvements in the last years have provided us with a wealth of knowledge about the dominant types of microorganisms in the oceans and their activities, there are still many research questions that have not been solved, either because the appropriate methodologies have not yet been developed or applied or, more often, because answering these questions implies an amount of work and resources that is out of the reach of most laboratories. In this position paper, the expert working group identified a set of key societal and scientific questions (Executive Summary Box C) and provided high-level recommendations on key future research priorities and needs (Executive Summary Box D) and proposed an approach with solutions to addressing these questions (Executive Summary Box E).

Executive Summary Box B.
Major achievements of marine microbiology in the last decades

- **1977**: Development of techniques to enumerate microbes in the ocean reveals, for the first time, the vast microbial abundance;
- **1979**: Microbes discovered in hydrothermal vents;
- **1980**: *Bacteria* in the ocean are shown to be actively synthesizing DNA and RNA, growing and reproducing with similar growth rates to those of algae and with production rates about one third of those of algae;
- **1982**: Discovery that *Bacteria* are actively predated by a group of highly specialized small protists, the heterotrophic nanoflagellates;
- **1983**: It is established that most primary production in the ocean is carried out by microbes smaller than 2 µm;
- **1986-1988**: Research shows that most predators of *Bacteria* and *Algae* are indeed mixotrophic flagellates and ciliates;
- **1989**: Use of flow cytometry allows the discovery of the picocyanobacteria, *Prochlorococcus*, the most abundant photosynthetic organisms on Earth;
- **1990**: First cultivation-independent assessment of marine bacterial diversity through rRNA analysis shows that they are highly diverse and that most groups in the ocean were previously unknown;
- **1990**: First phylogenetic stains used for marine plankton;
- **1989-1990**: The discovery of the vast numbers of viruses (ten million per millilitre) in the ocean and their role in recycling of nutrients, keeping in check successful microorganisms and horizontal gene transfer, all of which generates and maintains microbial diversity;
- **1992-1994**: High abundances of *Archaea* in marine plankton, even in cold, well-oxygenated waters;
- **1997**: Bacterial activity in the ocean is found to be so important that bacterial respiration in the oligotrophic ocean is higher than primary production;
- **2000**: Discovery of photoheterotrophy in the sea by metagenomic techniques (proteorhodopsine) and by other techniques (AAP bacteria);
- **2001**: Scientists find that a large unknown diversity also exists among the smallest protists;
- **2002**: Isolation of SAR11 in pure culture. This is followed by the discovery and isolation of many new marine microorganisms and new metabolic pathways (*Pelagibacter*, the most abundant bacterium in the ocean; temperate *Archaea* such as the ammonium-oxidizing *Thaumarchaeota*; anaerobic ammonium oxidizing (anammox) *Planctomycetes*; unicellular N_2-fixing *Cyanobacteria*), etc.
- **2003**: Genomes of environmentally important microorganisms are sequenced (e.g. *Cyanobacteria*)
- **2004**: A metagenome is produced from the Sargasso Sea revealing a large amount of unknown microbial functions in the ocean
- **2006-2007**: High-throughput sequencing technologies introduced in marine microbial ecology reveal that bacterial diversity is larger than expected and that the ocean contains an extremely large number of microbial genes, most of them of unknown function.

Executive Summary Box C.
Key scientific and societal questions related to marine microbial diversity research

Key societal questions:

1. What is the role of marine microbial communities in relation to major societal challenges such as global change, the sustainable supply of healthy food and energy, human health and environmental health?
2. If we have a better understanding of the role of marine microbial communities for the above functions, is it possible to manage them (better)? If so, how?
3. What opportunities can marine microorganisms provide for innovation in support of the European bioeconomy?

High level **scientific questions** that will need to be addressed:

1. What is the nature of microbial diversity in the oceans?
2. Which taxa are most relevant in terms of function across different ocean ecosystems?
3. What metabolic pathways exist?
4. What regulatory and signalling networks exist?
5. How do environmental factors influence metabolism and regulation within and between taxa?
6. How do microbial interactions influence ecology and ecosystem functioning?

Executive Summary Box D.
Strategic research priorities and needs

I. Ensure appropriate and accurate data acquisition to obtain more and better data;
II. Build a registry for samples and genetic materials;
III. Develop innovative cultivation approaches and, in turn, establish model organism databases and genetic systems for marine organisms;
IV. Improve classical methods/approaches and develop novel techniques and combine their results with abiotic and biotic information. By applying such an approach it will become possible to generate a more precise picture of the function, interaction and diversity within marine microbial food-webs;
V. Generate a dense network of data on the marine ecosystem by increasing both the number and the technological capacities of marine observatories;
VI. Contextualise sequence data with a set of standardized parameters (metadata) of habitats;
VII. Improve data management to be able to coordinate sampling, sample analyses and processing of data and make this data openly available for users;
VIII. Ensure interoperability of multidisciplinary data repositories;
IX. Implement well-curated and integrated ecological/diversity/and genetic databases;
X. Fully implement a systems approach to microbial ecology to understand ecosystem functioning;
XI. Prepare European researchers to be able to participate in, and benefit from, data delivered by international projects and programmes.

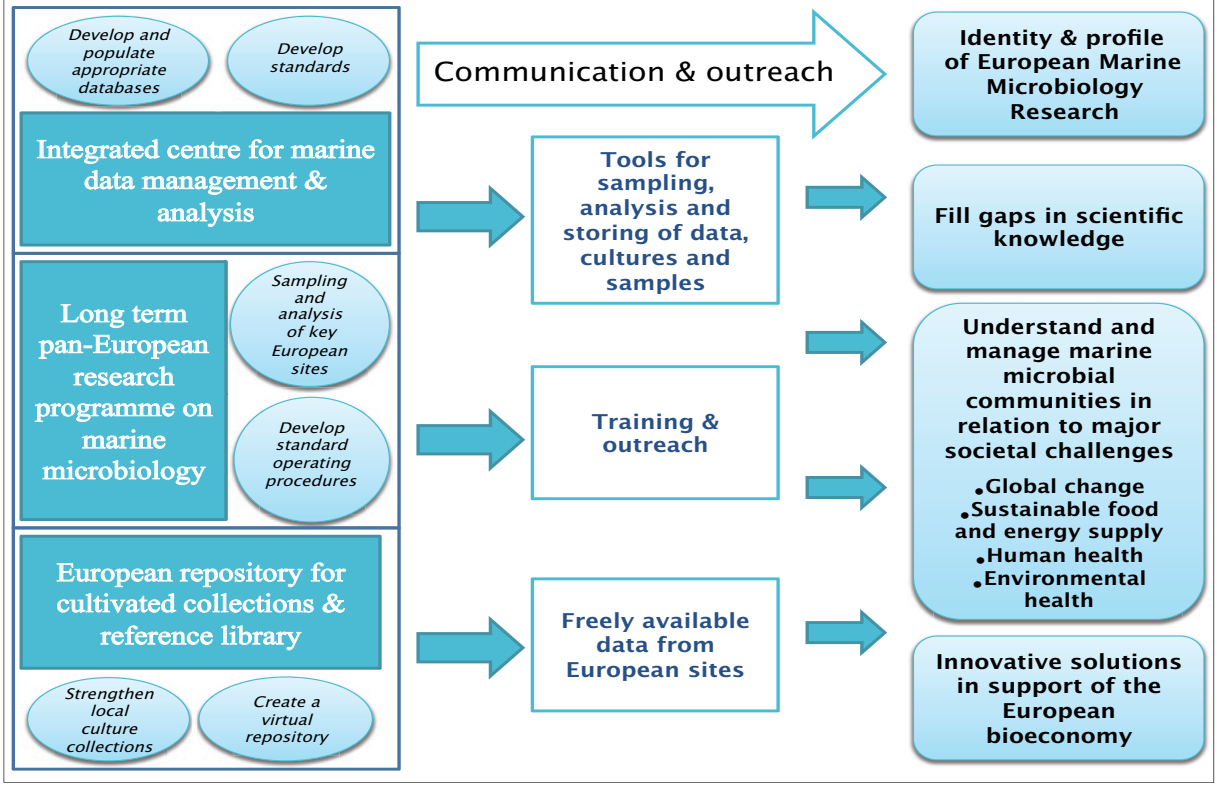
The high-level recommendations are:

- **Recommendation 1:**
  Establish a coordinated pan-European research programme on marine microbiology
- **Recommendation 2:**
  Create a European repository for cultivated collections and a reference library
- **Recommendation 3:**
  Create an integrated, multidisciplinary European Centre for marine data management and analysis
- **Recommendation 4:**
  Promote interest in marine microbial research and improve training and education

These four high-level recommendations are interdependent, complimentary and positively reinforcing. Creating better visibility for the importance of marine microbial diversity and associated research will assist in creating the necessary leverage for appropriate science policy measures and funding of relevant capacities and research activities. At the same time, the realization of a major pan-European research programme on marine microbiology will create a wealth of material and opportunities for outreach and education which will undoubtedly raise the profile of marine microbial research in Europe.

A multidisciplinary European Centre for marine data management and analysis is necessary to prevent the large-scale implementation of omics technologies turning into a disruptive process. It will provide critical support and capacity to deal with the wealth of samples and genetic information that will come on-stream through the research in the framework of the pan-European research programme on marine microbiology. Likewise, material gathered during sampling campaigns would be preserved and made available to a larger community if the proposed European repository for cultivated collections and a reference library becomes reality. The combined set of recommendations, when realized, will provide an enormous positive impact on European marine microbial research and provide the necessary push to improve Europe's competitiveness in this important research area (Executive Summary Box E)

Executive Summary Box E.
Schematic overview of main recommendations and their expected impacts in relation to addressing the key societal and scientific challenges



1. Introduction

Microorganisms are defined by their size: any organism that is too small to be observed in sufficient detail by the unaided human eye is a microorganism. This includes basically any organism smaller than 0.1 mm. The smallest well-known marine microorganism *Candidatus* Pelagibacter ubique HTCC1062 measures only 0.5×0.15 µm and hence the size range of microorganisms comprises more than three orders of magnitude, more or less the same as for macroorganisms. All three domains of life (Figure 1.1) comprise microorganisms while *Bacteria* and *Archaea* are comprised exclusively of microorganisms. All macroorganisms are *Eukarya* but the vast majority of the eukaryotic phylogenetic groups are nevertheless also microorganisms.

Bacteria and *Archaea* are unicellular organisms with cell shapes basically varying from spherical to elongated rods although a variety of other morphological characteristics are also possible. Some of these microorganisms may produce multicellular forms and a few are even capable of a simple form of cell differentiation. Sometimes microscopic organisms produce macroscopic appearances such as microbial mats and water blooms. Microorganisms can also produce multicellular structures that develop their own behaviour through cell-to-cell communication and a distribution of tasks.

Although *Bacteria* and *Archaea* do not differ much in appearance, they show an enormous range of metabolic capacities. *Eukarya, on the one hand,* are limited in their metabolic capacity to aerobic respiration of organic carbon (animals, protists) and to oxygenic photosynthesis connected to the fixation of CO_2 (plants, protists), and basically cycle carbon between inorganic and organic forms producing and consuming oxygen. The other two domains, on the other hand, take care of the full plethora of energy generating reactions that are possible by moving electrons from oxidized and reduced forms of e.g. iron, nitrogen, sulfur, etc., including the aerobic respiration and oxygenic photosynthesis that are characteristic of *Eukarya*. *Bacteria* and *Archaea* therefore are exclusively driving all other biogeochemical cycles while at the same time taking an important part in the global carbon and oxygen cycles.

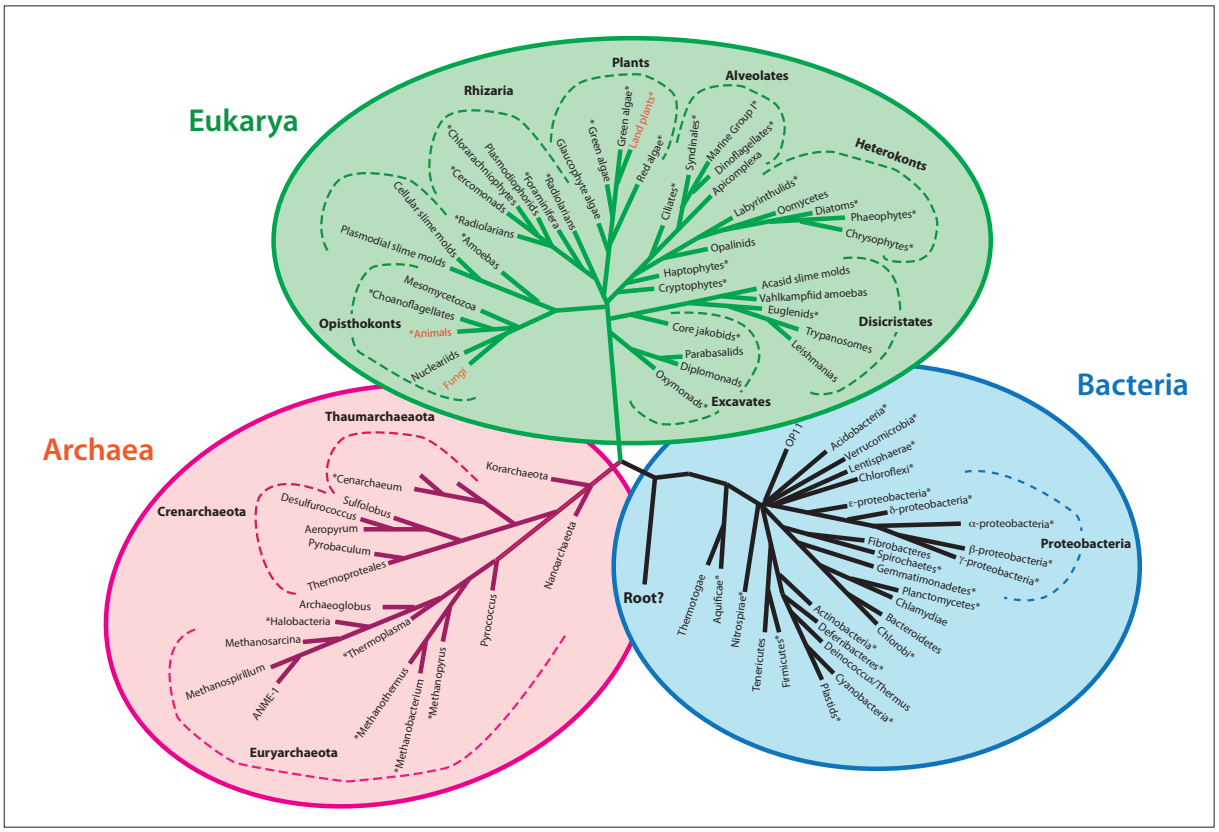


Figure 1.1. A tree of life including *Archaea, Bacteria* and *Eukarya.* The groups in black colour are mostly or completely microbial. The groups in red are not. Groups with asterisks are marine, or include a large amount of marine organisms. (Figure adapted from Baldauf 2003 using the colouring scheme of Barton *et al.* 2007 and the archaeal groups following Brochier-Armanet *et al.* 2008).

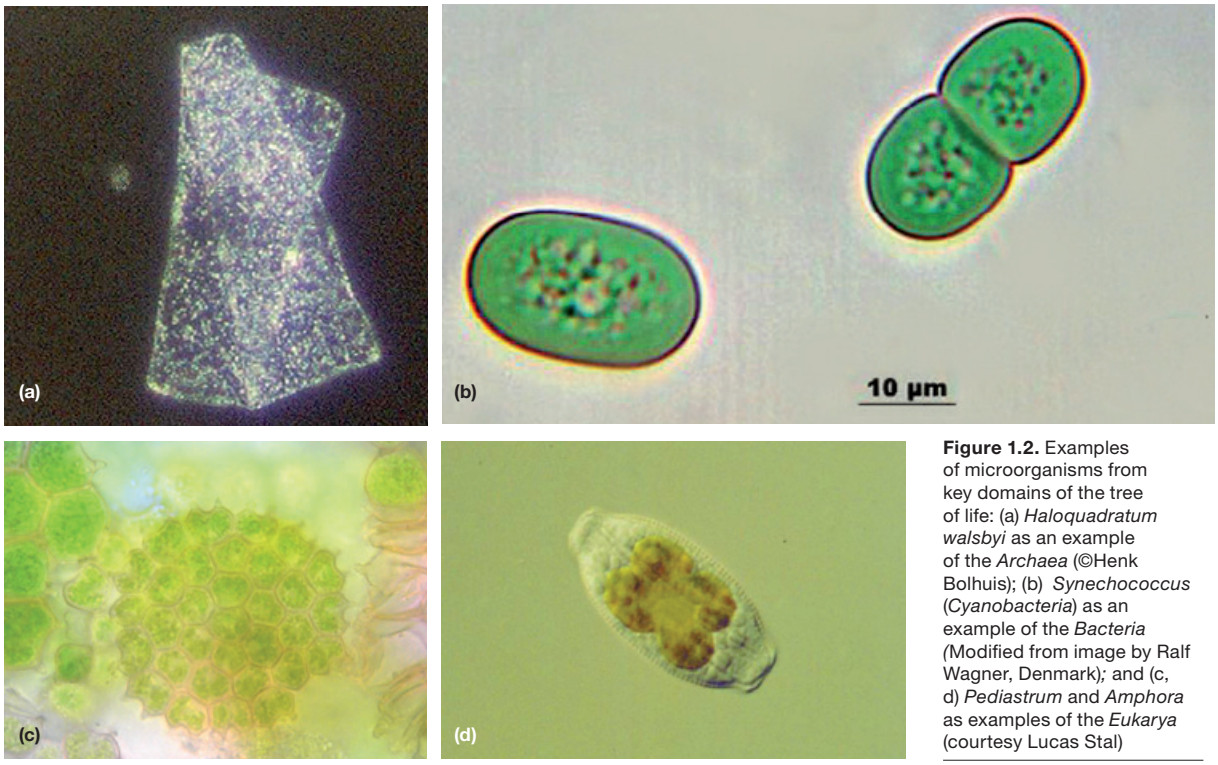


Figure 1.2. Examples of microorganisms from key domains of the tree of life: (a) *Haloquadratum walsbyi* as an example of the *Archaea* (©Henk Bolhuis); (b) *Synechococcus* (*Cyanobacteria*) as an example of the *Bacteria (*Modified from image by Ralf Wagner, Denmark*);* and (c, d) *Pediastrum* and *Amphora* as examples of the *Eukarya* (courtesy Lucas Stal)

What about viruses?

Organisms are defined by their capacity of independent cell division and therefore possess the full machinery for the duplication of DNA and for the binary division of the cell. Viruses, in contrast, require a living organism in order to duplicate their genetic material and to synthesize the virus particle. They do not, therefore, belong to any of the three domains of life. Nevertheless, viruses are of prime importance for generating and maintaining diversity amongst living organisms in all three domains of life through the mediation of horizontal gene transfer. They are also critical for the functioning and balance of the microbial food web and biogeochemical cycles by facilitating the production of nutrients for growth and maintaining high diversity, all by preventing the dominance of the most successful microorganisms. Viruses not only occur in numbers exceeding an order of magnitude those of microorganisms but also represent an untold genetic diversity. In all microbial ecosystems, viruses represent a considerable amount of the particulate organic matter. For these reasons we will consider the viruses as part of the microbial world and discuss them as such in this position paper.

What are marine microorganisms?

In general, the distinction between terrestrial microorganisms and aquatic microorganisms is rather clear, even though this is attributed in part to the traditional and historical differences in the scientific approaches between aquatic and soil microbiology. The distinction between freshwater and marine microbiology is less clear as it is based on the assumption that salinity is a factor discriminating aquatic microbial life, i.e. microorganisms growing in freshwater do not occur in the ocean and vice-versa. In fact this is not the case. Certain microorganisms occur in a wide range of aquatic environments with different levels of salinity.

For the purpose of this position paper, we will consider any microorganism that grows in a marine environment as a marine microorganism. This includes low-salinity environments adjacent to the sea such as estuaries and brackish-water environments (e.g. Baltic Sea) which are traditionally considered “marine” and hypersaline environments such as deep sea brines and solar salterns receiving seawater.

Although traditionally not considered marine, what is stated in this document will in many cases be relevant to inland saline and non-saline waters as well. This is because, in general, the scientific approaches in fresh-

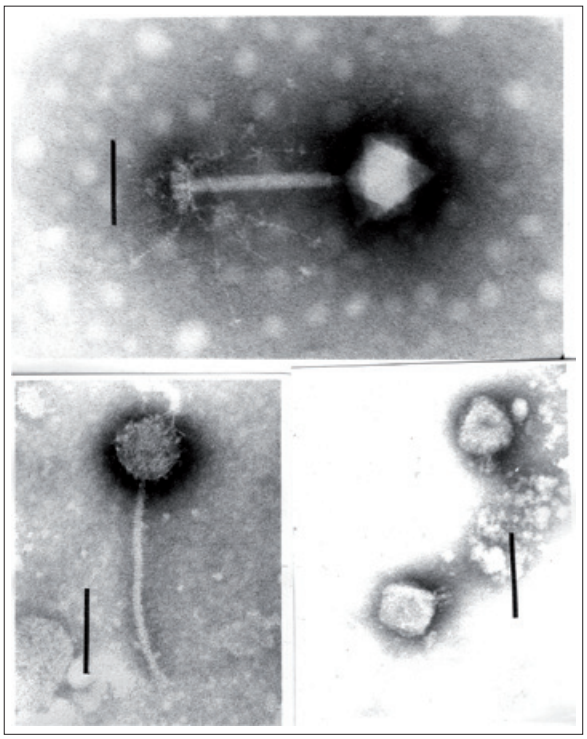


Figure 1.3. TEM micrographs of representatives of the most common phage families in the surface ocean: *Myoviridae,* represented by the cyanophage S-PM2 (above), *Sipoviridae*, represented by the bacteriophage VP6 (below left) and *Podoviridae*, represented by the bacteriophage H100/1 (below right). Scale bar 100 nm. (©Professor Hans-Wolfgang Achermann, Laval University, Canada; from Sandaa, R-A. 2009).

water and marine aquatic microbiology do not differ fundamentally, except for matters of scale: oceans and seas cover a much wider area than freshwater bodies; the depth (and hence pressure) is an order of magnitude larger in the ocean; and there is basically one global ocean that is connected through a complex system of currents.

Microorganisms are those organisms that are too small to be observed in sufficient detail by the unaided human eye which includes roughly any organism smaller than 0.1 mm. For the purpose of this position paper, we consider **marine microorganisms** those microorganisms that can grow in marine environments, independent of whether they are more abundant in other aquatic or terrestrial environments. Although traditionally not considered marine, what is stated in this document will in many cases also be relevant to inland saline and non-saline waters.

What is the problem?

Marine microorganisms are of critical importance to the health of our environment and our well-being. They are integral to all major biogeochemical cycles, fluxes and processes occurring in marine systems where elements move between oxidised and reduced forms. Microbes are extremely abundant and diverse and produce and release carbon products that are key in the regulation of the Earth's climate, particularly CO_2 and CH_4. Marine microorganisms also provide essential goods and services to our society in terms of production of oxygen, supporting sustainable supply of food, regulating the health of the marine environment and providing an largely untapped source of genetic information and biomolecules for use in industrial and medical applications and products.

In spite of their importance, very little is known of marine microbial diversity, how many types of microorganisms are present in the oceans, and what the role of each of them is – i.e. their ecological functions and interactions. In fact, microbiologists are still uncertain about how to define a microbial species. Moreover, the vast majority (90-99%) of marine microorganisms cannot be cultured under standard laboratory conditions and as such their growth behaviour and physiology cannot be studied, an approach that has proven so successful throughout the 20th century in the case of medically-important microorganisms.

It was only with the development of a molecular toolbox to detect and sequence DNA from the natural environment that information about the exceptional prokaryotic diversity in the oceans began to accumulate. An example of these molecular tools is the powerful Polymerase Chain Reaction (PCR)-based methods that have been established for direct amplification, cloning, and analysis of ribosomal RNA (rRNA) genes from the environment. Over the last 30 years two million rRNA gene sequences have been added to public repositories and represent an impressive hallmark of this development. The full cycle rRNA approach makes use of these data for the design of specific probes to label microorganisms using fluorescent dyes. With the development of this Fluorescent In Situ Hybridisation (FISH) technique it was, for the first time, possible to visualize and enumerate specific microbes by preserving the microbial communities in their natural habitats. The advent of the rRNA gene sequence as a gold standard in analysing microbial diversity and abundances has paved the way for microbiology to answer the **first of three basic ecological questions: "What is out there?**

While fundamental microbial functions have been discovered through traditional approaches since studies of Antony van Leeuwenhoek in the late 17th century,

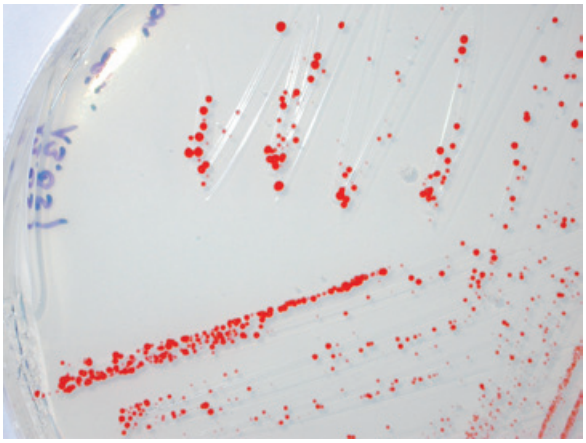


Figure 1.4. Example of a microorganism culture on a laboratory Petri dish. However, the vast majority of known marine microorganisms cannot be cultured under standard laboratory conditions. (Courtesy Frank Oliver Glöckner)

more recently, marine microbial genomics ranging from the study of the genomes of model organisms to the wealth of meta-omics approaches (e.g. metagenomics, metatranscriptomics and metaproteomics) has proven to be very successful **to target the second basic ecological question "What are they doing?"** (see also Chapter 3). The cultivation independent approaches, in particular, have resulted in an explosion of information on marine microbes. For example, the first part of the Global Ocean Sampling (GOS) project, which sampled the North Atlantic, Caribbean and a small part of the Pacific Ocean, added DNA sequence information that was equivalent to 50% of all protein-encoding sequences that had previously been deposited in GenBank. The GOS project confirmed that marine microbes are diverse, revealing how little is known about the genetic information of natural assemblages. At the same time, this study also highlighted the difficulties of making sense of metagenomic sequence data: a significant proportion of the open reading frames (ORFs, which are presumed to equate to genes) could not be characterized because there were no similar sequences in the databases. This demonstrates that even billions of nucleotides do not automatically provide much information when it comes to the understanding the function of organism or a community.

Information Box 1. Basic questions of marine microbial diversity research

While recent technological developments and scientific discoveries have been substantial, we still lack a major understanding at all levels of the basic ecological questions in relation to the microorganisms in our seas and oceans. These fundamental questions are:
1. What is out there? Which microorganisms are present in our seas and oceans and in what numbers do they occur?
2. What are they doing? What functions do each of these microorganisms perform in the marine environment and how do they contribute to the global cycles of energy and matter?
3. What are the factors that determine the presence or absence of a microorganism and how do they influence biodiversity and function and vice versa?

This lack of understanding limits our ability to assess the precise relationship between the diversity and function of marine microorganisms on the one hand, and anthropogenic global changes on the other. For example, currently we are neither able to answer the question of how climate change influences the diversity and function of microbes nor can we determine the exact role and contribution of marine microorganisms in moderating climatic changes. This lack of knowledge also limits our ability to tap the full potential of the marine realm to develop goods and services using marine microorganisms.

More cultures of environmentally relevant marine *Bacteria*, *Archaea* and viruses can backup these efforts. Most culture collections are based on readily cultivated microbes. When these organisms were isolated, there were no techniques to establish if the isolate was abundant in the natural environment or even if it had any relevant function. Molecular biology has changed that, and the isolation of new cultivable microbes can now be based on their abundance and relevance in defined marine habitats.

At the same time, there are also a number of novel and innovative approaches to the isolation of new potential model microorganisms. For example, in 2002 Rappé used a dilution approach to isolate one SAR11 representative, *Candidatus* Pelagibacter ubique, the bacterium which can be found in almost all marine biodiversity or metagenomic studies. Hence, methods do exist for isolating useful model microbes from the natural environment, but they remain labour-intensive. Nevertheless, these methods are probably the only way in which relevant bacteria can be brought into culture. Having the relevant cultures in hand, the whole toolbox

of genomics, genetics, biochemistry and physiology can be applied to start unravelling the function of hypothetical genes.

To address the functional potential of organisms, integrated approaches are needed that take into account the results of the **third basic ecological question "What are the environmental factors of relevance and how do they influence biodiversity and function?**

Full geo-referencing has the potential to be extremely useful in this respect, especially for the open ocean, where any kind of genomic data can be easily linked with measured and remote sensing data based on location, time and depth. Subsequently, the diversity and abundances of organisms, the genomic information, as well as gene expression information at the transcriptome and proteome level can easily be integrated with measured or interpolated environmental parameters. Unfortunately, these so called contextual data or metadata of sampling campaigns are often incomplete and stored in hand-written laboratory notebooks or on individual computers. What is finally published represents only a small subset of the original data and is often only "human readable" and cannot be assimilated into databases without major efforts in text recognition.

The lack of knowledge on the organism and habitat level greatly hampers any kind of Europe-wide or even international medium to large scale data integration. To move towards a better understanding of the marine ecosystem as a whole, community efforts to improve data management and standardization are a prerequisite.

The way forward

Over the past years, considerable progress has been made to advance marine microbial research in Europe through better coordination and collaboration of research activities, strategies and infrastructures. Major pan-European and macro-regional initiatives and projects such as the marine Networks of Excellence on marine biodiversity (MarBEF) and marine genomics (MGE) have contributed to the initial development of a European Research Area (ERA) for marine sciences and developed a strong momentum to further strengthen

Information Box 2. This Position Paper: objectives, underlying questions and approach

This publication is the result of the endeavours of a Marine Board Working Group of experts set up to highlight the importance of marine microbial biodiversity research and put forward ways to address the current lack of knowledge and barriers to progress. To this end, the Working Group general objectives were to:

- Re-emphasise the importance and role of microbes in the marine environment;
- Identify strategic areas for Europe in the realm of marine microbial diversity research, with specific reference to ecosystem functioning, biogeochemical cycling and environmental change;
- Provide recommendations and a roadmap to guide European research in the medium term (to 2020) to substantially increase our knowledge of marine microbial diversity and its role in, and response to, environmental change and to improve the competitiveness of European research in this field;

To address the above objectives, the Working Group formulated key underlying societal and scientific questions which will need to be tackled in future research in the area of marine microbiology.
The three **societal questions** are:

1. What is the role of marine microbial communities in relation to major societal questions such as global change, the sustainable supply of healthy food and energy, human health and environmental health?
2. If we have a better understanding of the role of marine microbial communities for the above functions and are able to measure this, is it possible to manage it (better)? If so, how?
3. What opportunities can marine microorganisms provide for innovation in support of the European bioeconomy?

The high level **scientific questions** that will need to be addressed are:

1. What is the nature of microbial diversity in the oceans?
2. Which taxa are most relevant in terms of function across different ocean ecosystems?
3. What metabolic pathways exist?
4. What regulatory and signalling networks exist?
5. How do environmental factors influence metabolism and regulation within and between taxa?
6. How do microbial interactions influence ecology and ecosystem functioning?

To answer the above societal and scientific questions, this position paper assesses the current state of the art in terms of scientific knowledge and research capacities available in Europe. It identifies what gaps exist in terms of knowledge and available research frameworks and highlights essential future research priorities and requirements to substantially increase our ability to address these issues.

1. Introduction

European capacity on marine microbial research (see also Chapter 5). Nevertheless, major gaps in our scientific knowledge still exist and marine microbial research remains largely uncoordinated and fragmented with the majority of funding coming from national resources.

To deal with these issues and address the identified societal and scientific questions (see Information Box 2), this position paper assesses the current state of the art in terms of scientific knowledge and research capacities available in Europe. It identifies what knowledge gaps exist and highlights essential future research priorities and requirements to substantially increase our ability to address these issues. To this end, the position paper starts with an overview of what we know, what we don't know, and what we should know about marine microbial diversity which reveals the important role, impact and societal benefits provided by marine microorganisms (Chapter 2). Next it considers the most important existing and emerging technologies and research toolkits which have shaped marine microbial research and will continue to drive it forward (Chapter 3). Available resources and gaps in marine microbiology observation and data infrastructures are considered in Chapter 4. Finally, Chapter 5 reflects on the European status of marine microbial diversity research and considers the main research gaps which need to be addressed to answer the basic scientific and societal questions and provides science policy recommendations to bridge the gaps.

This paper is intended in first instance for those who determine and develop research agendas, the Research Funding Organisations (RFOs), programme managers and science policy advisors and developers both at the national and European level. At the same time, the expected outcome is also intended to strengthen the marine microbial research community by stimulating networking and developing common views of expert scientists and the major marine science organisations in Europe.

2. Marine microbial diversity: current knowledge and gaps in our understanding

2.1 Introduction

Life on Earth was microbial for more than 3.2 billion years and it developed and maintained every biogeochemical cycle that exists today. Microorganisms evolved the plant-like photosynthesis that resulted in the enrichment of the biosphere and atmosphere with oxygen. Microbes shaped the Earth and made it habitable for macroscopic life including humans. Microbes continuously changed the environmental conditions on Earth and have adapted to global environmental changes as they happened.

The oceans comprise the largest continuous ecosystem on Earth. Marine microbial communities are an integrated part of the ocean and are responsible for the uptake of a large part of the carbon dioxide that human society emits into the atmosphere and that causes global warming. The biological pump transports CO_2 to the seafloor, a process entirely driven by microbes. The constantly increasing level of atmospheric CO_2 is causing the acidification of surface waters of the ocean. This may lead to the dissolution of carbonates, and the change of the carbonate equilibrium could eventually lead to an ocean that would release CO_2 into the atmosphere instead of taking it up. Marine microbial communities respond to these changes in hitherto unpredictable ways.

Marine microbial communities are the basis of the ocean food web and, hence, produce the food for all life in the ocean. Fisheries are supplying a major part of the protein demand of human society and they depend directly from adequately functioning microbial communities. This awareness has resulted in a growing interest in sea-based aquaculture for sustainable production of sea-food, as well as for the efficient recycling of nutrients in these production plants ensuring a minimum of negative effects on the environment.

More recently, it has been realized that marine microorganisms may be relevant agents for the sustainable production of energy. Sunlight at the ocean surface is used by *Cyanobacteria* and microalgae to produce biomass by fixing CO_2 while producing O_2. Fossil fuels such as natural gas and oil were produced by the same organisms and accumulated in the Earth's crust for hundreds of millions of years. Attempts are now being made to apply this process directly so that the CO_2 produced from burning fuels could be moved from the atmosphere to be fixed back as renewable fuel.

Society relies on marine microbial communities for its own health as well as for the health of the environment. Marine microbial communities are the source of a large variety of bioactive compounds which may have medical applications and, as such, contribute to human health. Marine microbial communities also provide a variety of services, such as bioremediation of polluted environments.

2.2 Role and impact of marine microorganisms

2.2.1 Introduction

The environments in which marine microorganisms are found are diverse. First of all, many microorganisms live as plankton in the water column, while other types live a benthic lifestyle on the seafloor as biofilms or as microbial mats (Figure 2.1).

A virtually unexplored marine environment is the deep subsurface, the Earth material near but not exposed to the sea bottom that can be found at about 50m below the surface of the Earth's crust, and which extends variably downward, up to 2.8km. Furthermore, we should consider specific environments such as deep sea hot- and cold hydrothermal vents, microorganisms that live attached to marine plants (epiphytes), or animals, and in fact on any submersed marine surface (including fouling of ship hulls). Last but not least we must consider microorganisms that live inside and in symbiotic relationships with marine organisms. Sponges and corals are well-known examples of organisms which have symbiotic relationships with microorganisms in the marine environment, but any fish in the sea carries its own specific microbiota.

As a rule of thumb we may say that ocean water contains one million microorganisms per millilitre. This number may be an order of magnitude higher in coastal waters which have higher productivities and loads of organic matter and nutrients. Likewise, the deep ocean may contain one to two orders of magnitude less microor-

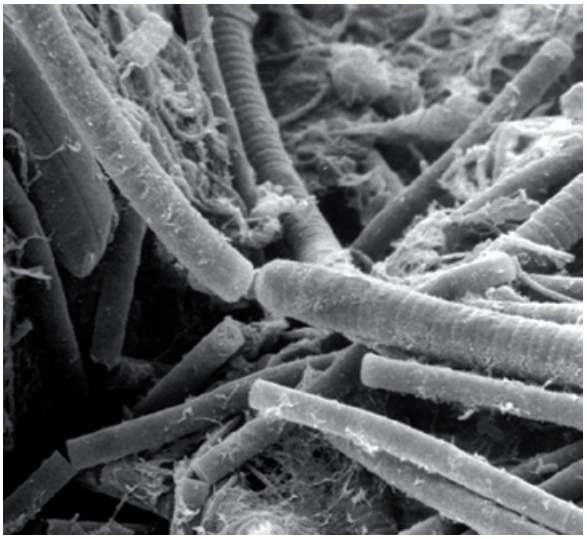


Figure 2.1. Microbial mat composed of *Cyanobacteria* and other microorganisms (courtesy Lucas Stal)

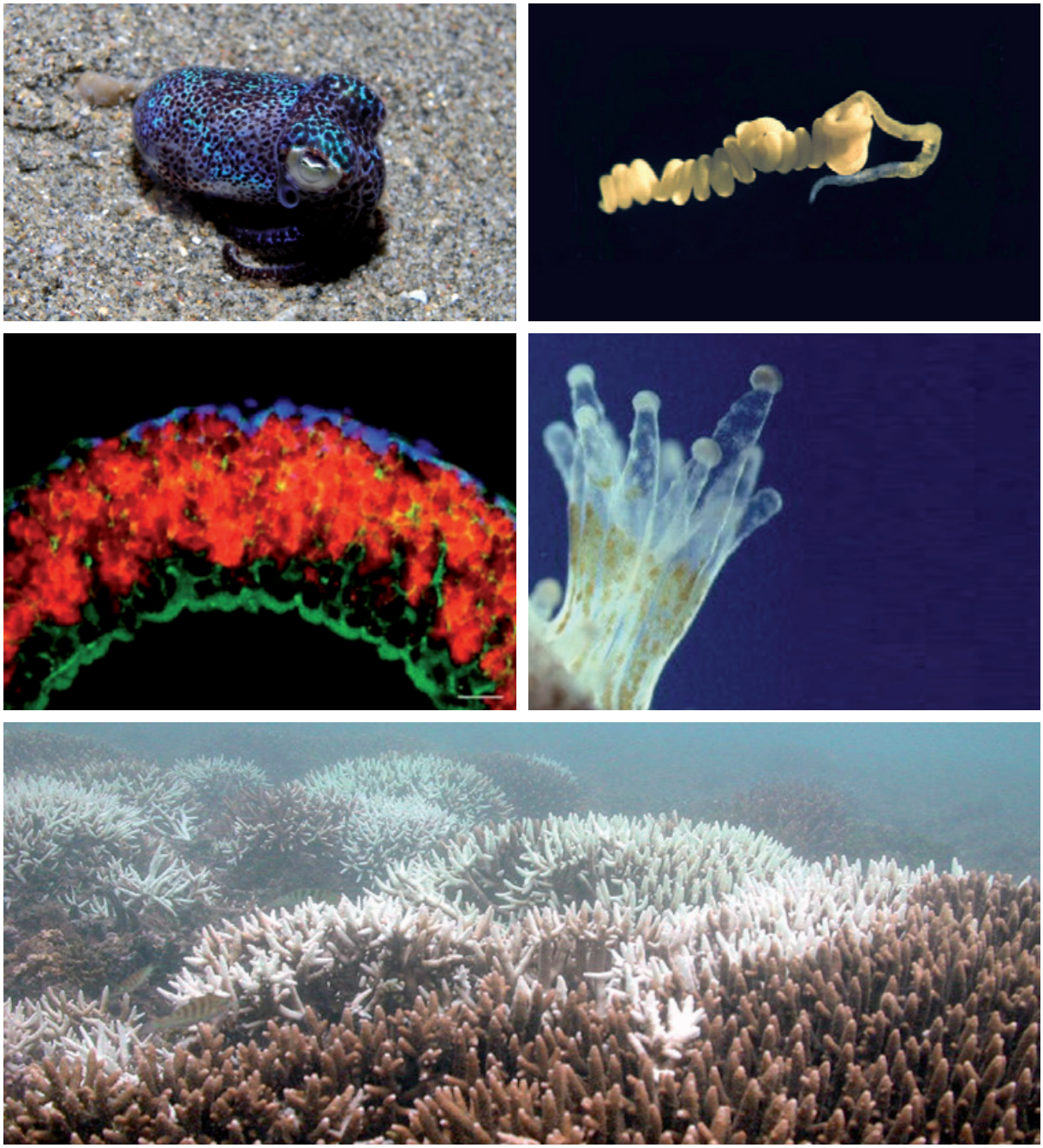


Figure 2.2. Examples of symbiotic marine microorganisms: Bobtail squid (*Euprymna berryi*) from East Timor with light organs filled with luminescent microorganisms (©Nick Hobgood) (above left); Gutless *oligochaetes* are a unique group of marine worms appearing bright white because of the elemental sulfur stored in their bacteria endosymbionts (above right). Epifluorescence image of the symbiont-containing region of the *O. crassitunicatus* worm's body wall (middle left) (Courtesy Anna Blazejak); Corals have small symbiotic single-cell algae residing in their tissue called *Zooxanthellae* (middle right) which photosynthesize in the presence of sunlight to make sugars which is given to the coral animal and may account for up to 90% of the coral animal's overall nutritional needs. *Zooxanthellae* give coral its coloration, with the specific color depending on the particular clade. Coral bleaching is the loss of intracellular endosymbionts through either expulsion or loss of algal pigmentation. Under stress (e.g. due to changes in temperature or occurrence of by pathogenic bacteria), corals may expel their *zooxanthellae*, which leads to a lighter or completely white appearance, hence the term "bleached", and ultimately the death of the coral (bottom). (©Bruce Monger, NOAA).

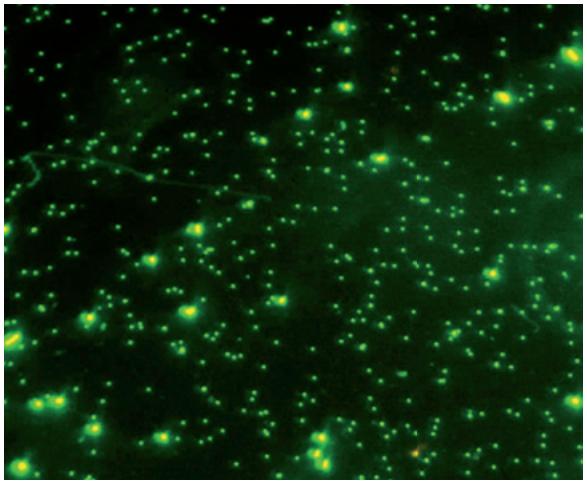


Figure 2.3. Epifluorescence micrograph of prokaryotes and viruses in a seawater sample stained with a fluorescent dye, SYBR Green I. The dye specifically stains doubled-stranded DNA (dsDNA). Smallest dots are viruses and larger ones are prokaryotes (*Bacteria* or *Archaea*). With about 1 billion bacterial cells and 10 billion viral particles per liter of seawater, bacteria and in particular viruses are by far the most common biological entities in the marine environment. (©Ruth-Ann Sandaa)

ganisms. Altogether, the number of microorganisms in aquatic habitats is estimated at 1.2×10^{29} and in the oceanic subsurface up to 3.5×10^{30}. The number of viruses in the ocean is estimated to be higher than 10^{30}. Very little is known about the diversity of marine microorganisms. The number of species of microorganisms has been estimated from as low as $10^4 - 10^5$ to as high as $10^6 - 10^7$, let alone that it is not always clear how to define a microbial 'species'.

2.2.2 Biogeochemical cycles

Oceans and coastal waters, including estuaries, harbour a tremendous diversity of *Bacteria*, *Archaea*, viruses, fungi, protists and microalgae able to transform C-, N-, P- and S-containing compounds in ways that influence their availability for biological production. Thus, the metabolism of marine microorganisms maintains the major biogeochemical cycles on Earth, including the significant production of oxygen required for aerobic life and the biological removal of carbon. The balance of all these cycles and compounds controls the dynamics of all ocean biomes. Thus, studying the ecology of marine microbial communities is essential for an understanding of ecosystem function.

Information Box 3. Biogeochemical cycles in the sea

A biogeochemical cycle is a pathway by which a chemical element or molecule moves through both biotic (biosphere) and abiotic (lithosphere, atmosphere, and hydrosphere) compartments of Earth. A cycle is a series of change which comes back to the starting point and which can be repeated. The term "biogeochemical" indicates that biological, geological and chemical factors are involved in the process. The circulation of water and chemical nutrients like carbon, oxygen, nitrogen, phosphorus, calcium, etc. through the biological and physical world is achieved through biogeochemical cycling. In effect, these elements are recycled, although in some cycles there may be places (called reservoirs) where the element is accumulated or held for a long period of time (such as the ocean or a lake for water). Elements, chemical compounds, and other forms of matter are passed from one organism to another and from one part of the biosphere to another through biogeochemical cycles.

In Figure 2.4, microorganisms in the box in the middle represent the particulate organic carbon matter (POC) that recycles key elements (see also Figure 2.10. on the microbial and viral loops). Microorganisms include *Cyanobacteria*, microalgae, *Bacteria*, *Archaea*, ciliates, flagellates, zooplankton, and viruses. For simplicity, only the cycles of C, N, P, Fe, O and H and a part of the sulfur cycle are depicted.

Sunlight is the main driver of the system through photosynthetic CO_2 fixation, although a considerable part of the CO_2, especially in oxygen minimum zones (OMZ) or hydrothermal vent systems, may be fixed chemosynthetically. Photosynthesis results in the production of oxygen from water and is used for respiration (re-oxidation of reduced carbon, nitrogen, iron, sulfur) to become water again. Oxygen and carbon dioxide exchange with the atmosphere.

The fixation of CO_2 results in the formation of particulate organic carbon (POC) and dissolved organic carbon (DOC) both of which are respired back to CO_2. Under anaerobic conditions, CO_2 may be reduced to methane (by methanogenic bacteria) which is oxidized back to CO_2 by methanotrophic bacteria. Iron cycles from oxidized ferric iron (Fe^{3+}) to reduced ferrous iron (Fe^{2+}) and back catalyzed by bacteria as well as by (abiotic) chemical processes. In seawater, virtually all iron is oxidized which is insoluble and therefore this essential element is thought to be unavailable and to limit growth of microorganisms, even though it is one of the most abundant elements. The

cycle of phosphorus is simple. Phosphate is taken up and incorporated in the cell and is released as dissolved organic phosphorus (DOP). Alkaline phosphatases cleave the DOP. Seawater contains a large amount of sulfate which can be assimilated by microorganisms and serves as the source of sulfur. Other organisms depend on organic sulfur, mainly in the form of protein. Some algae produce DMSP (dimethyl sulfoniopropionate) which, when liberated, can be transformed by some bacteria and other algae to DMS (dimethylsulfide). This gas, which is responsible for the "smell of the sea", escapes to the atmosphere where it may cause cloud albedo and thereby represents a negative climate feedback (see also Figure 2.13).

Nitrogen is quantitatively one of the most important elements. Most organisms depend on ammonium, nitrate or dissolved organic nitrogen (DON). Ammonium can be converted anaerobically to N_2 or aerobically to nitrate, which is denitrified to N_2. Incomplete denitrification may lead to the formation of nitrous oxide (N_2O) which is a strong greenhouse gas. Only a few bacteria are capable of fixing N_2 and use it as a source of nitrogen. All these elements (and many more) are required for the synthesis of structural cell material. Dead biomass sinks to the ocean floor where part of it is stored on the geological time scale away from the biosphere (this process is called "the biological pump").

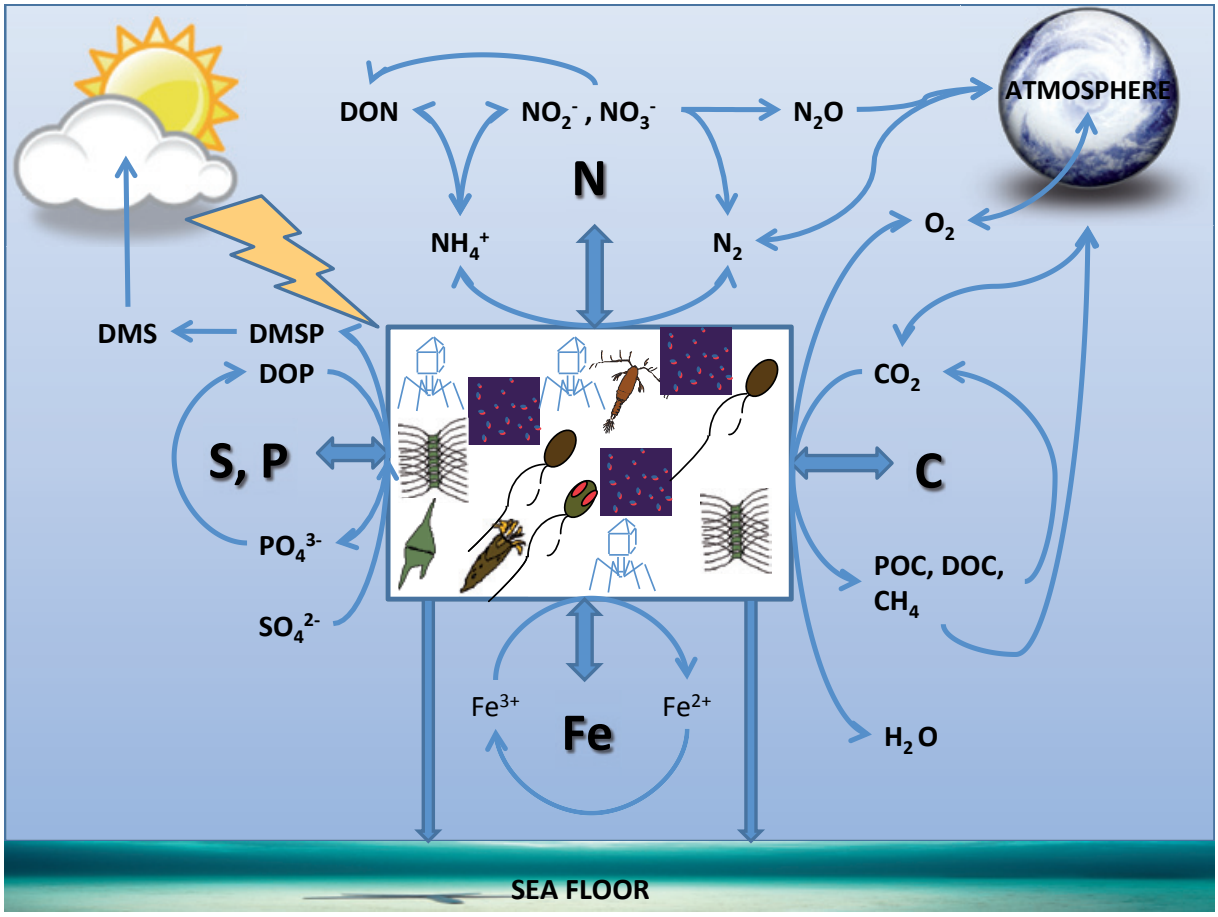


Figure 2.4. Biogeochemical cycles in the sea and the role of the microbial food webs (developed by Lucas Stal and JM Gasol)

2.2.2.1 Carbon

Inorganic carbon in seawater is in equilibrium with atmospheric CO_2. Carbon dioxide dissolves in the seawater where it is protonated to bicarbonate and carbonate that are in a pH dependent equilibrium. The fixation of CO_2 results in the shift of the carbonate equilibrium towards carbonate and consequently to an increase in pH. Atmospheric carbon dioxide dissolving in seawater will have the opposite effect on the pH. Calcium carbonate precipitation in seawater is in fact largely controlled by biology. Much of the calcium carbonate will sink to the ocean floor and be recycled over geological timescales. The increasing atmospheric CO_2 concentration causes a higher dissolution of carbonic acid and is expected to decrease the seawater pH (causing ocean acidification). This would also affect calcium carbonate precipitation or even cause dissolution of the massive amounts of calcium carbonate (stored in coral reefs, other carbonate platforms, shellfish, etc.).

The organic carbon produced by cyanobacterial and microalgal CO_2 fixation (and by sea grasses in shallow coastal areas) enters the food web and is liberated as particulate or dissolved organics through fecal pellets produced by grazers and predators, death of organisms, viral lysis and by various exudation processes. Carbon cycles between the most oxidized form (carbon dioxide) and the reduced forms: organic matter or in the most reduced form, methane. Seawater contains a remarkably high amount of dissolved or colloidal organic matter as well as particulate organic matter, which are all subject to (partial) degradation by the microbial community (*Bacteria* and *Archaea*). Part of the organic matter sinks to the ocean floor and the part that is not mineralized during that voyage is eventually buried and stored over geological timescales during which diagenetic processes transfer it to natural gas, petroleum, kerogen and bitumen. Together with limestone and other fixed carbonates, these pools represent by far the largest source of carbon on Earth. Anthropogenic activities speed up the recycling of a tiny part of these geological carbon deposits, but this is sufficient to cause a significant global change in the carbon cycle.

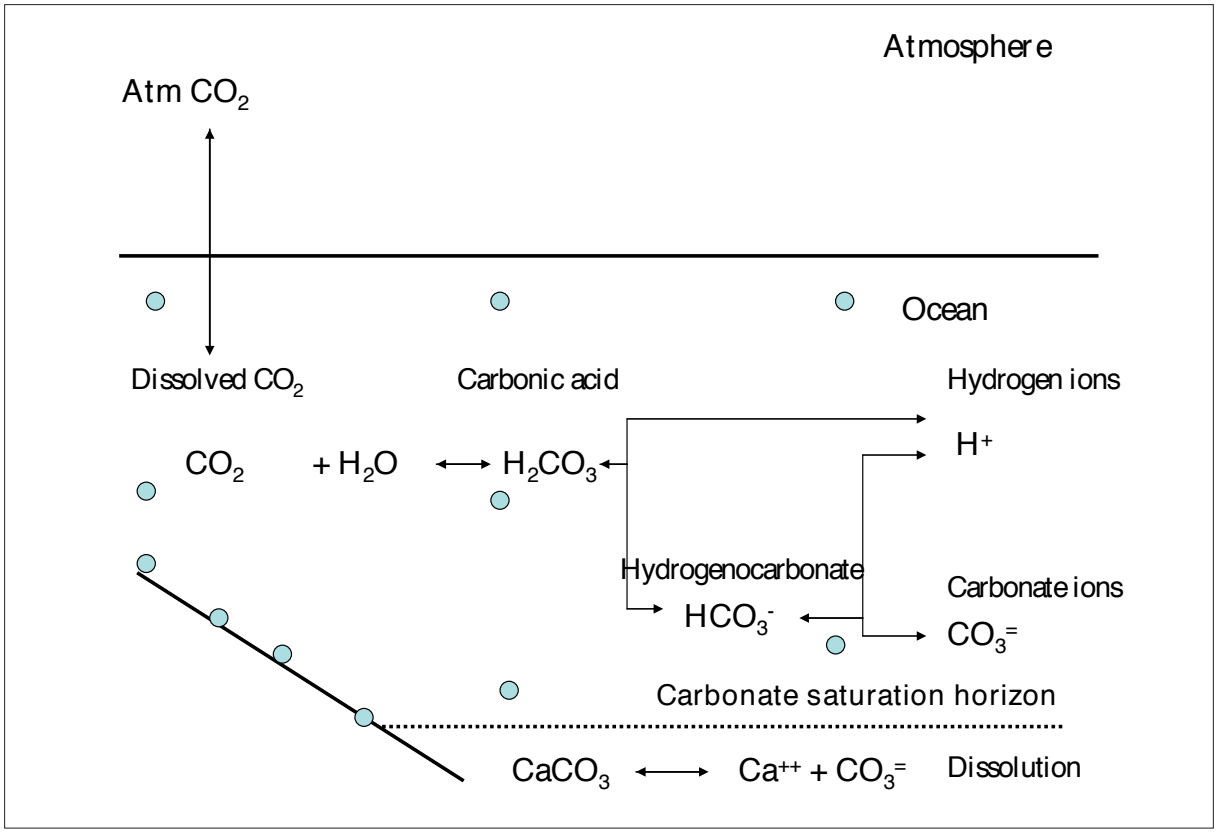


Figure 2.5. Chemical basis of the carbonate system in seawater. Calcareous shell-forming plants and animals (blue circles) can develop in waters which depths are shallower than the carbonate saturation horizon. Reversely, calcium carbonate dissolution is active in waters below this depth level.

2.2.2.2 Nitrogen

Nitrogen is an important element which may represent up to 10% of the dry mass of living organisms. Nitrogen occurs in amino acids (the building blocks of proteins), nucleic acids, bacterial cell walls and chlorophylls. Most organisms rely on combined forms as a source of nitrogen such as ammonium, nitrate, urea, or organic compounds (amino acids, proteins). A few specialized *Bacteria* and *Archaea* are capable of utilizing atmospheric dinitrogen (N_2), which is the most abundant form of nitrogen on Earth, while combined nitrogen is often in low supply. The simplest form of nitrogen cycling is the assimilation of any form of nitrogen for growth and the subsequent release of ammonia (deamination) after the death and mineralization of the organism. However, this mini nitrogen cycle is interrupted because of microorganisms that oxidize ammonium to nitrite. In the ocean this is done by *Thaumarchaeota*, a newly discovered group of *Archaea*. Other bacteria may oxidize this nitrite further to nitrate. Nitrate and nitrite are subject to denitrification, a bacterial process that converts these compounds to dinitrogen. Anammox (anaerobic ammonium oxidation), a newly discovered process, also converts ammonium and nitrite to dinitrogen. These processes are important in anaerobic marine sediments, in the ocean oxygen-minimum zones (OMZs) and also in marine snow, faecal pellets and similar aggregates in the water column. Finally, nitrite and nitrate can also be reduced to ammonia in a dissimilatory process (dissimilatory nitrite or nitrate reduction to ammonia, DNRA). Thus, the nitrogen cycle is entirely microbially driven. While most of the nitrogen is cycled rapidly by these biological processes, part of the nitrogen is buried and stored in the sediment (see also Figure 2.13).

2.2.2.3 Sulfur

With 28mM sulfate in seawater, the marine environment is dominated by sulfur, which is a micronutrient, essential as a component of a few amino acids and in iron-sulfur clusters in enzymes. A simple sulfur cycle is one that just includes uptake, assimilation and its liberation after death and mineralization of organisms. Quantitatively

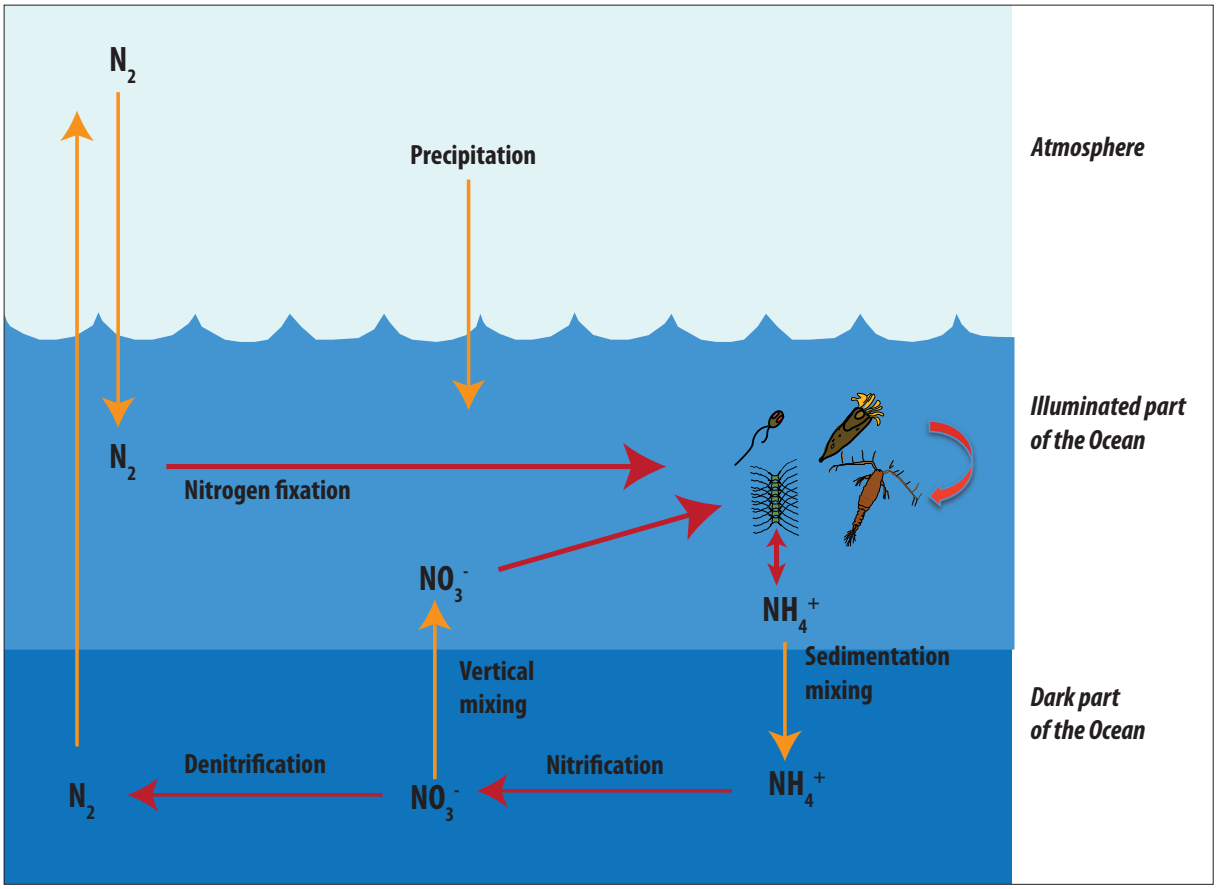
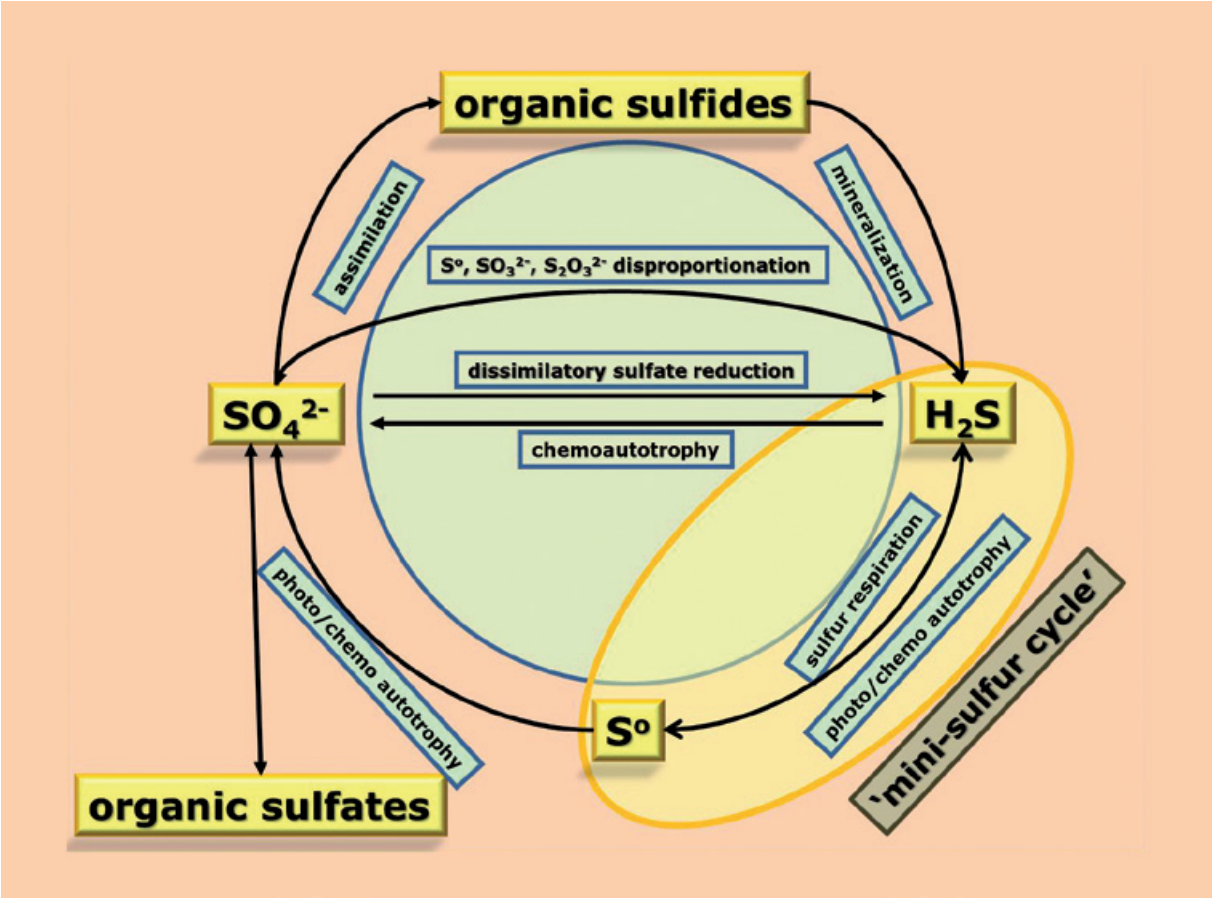


Figure 2.6. Schematic overview of the nitrogen cycle in the ocean (courtesy JM Gasol)



more important is dissimilatory sulfur metabolism. Sulfate (and other oxidized sulfur compounds) can be reduced to sulfide by anaerobic sulfate-reducing bacteria. Sulfide and other reduced sulfur compounds can be used as an electron donor for anoxygenic photosynthesis (particularly sulfide and elemental sulfur). Sulfide can be oxidized aerobically with O_2 or anaerobically with nitrate as the terminal electron acceptor. Because sulfide also reacts chemically with O_2, the biological process occurs usually at the aerobic-anaerobic interface in ecosystems. Anoxygenic phototrophic bacteria often produce elemental sulfur, which can be reduced back to sulfide as an electron acceptor in anaerobic respiration or as an electron sink in fermentation (sulfur reducing bacteria). The oxidation of sulfide to elemental sulfur and back to sulfide occurs rapidly in anaerobic ecosystems and is also known as the "mini sulfur cycle". Another set of microbial transformations of sulfur compounds is known as disproportionation or as sulfur fermentation.

Figure 2.7. Schematic overview of the the Sulfur Cycle. Sulfur occurs in redox states of -2 (as in H_2S, sulfide) and +6 (as in SO_4^{2-}, sulfate) and in any state in between. Transitions from one redox state into the other are catalyzed by microorganisms or in some cases also by chemical reactions. Altogether these transitions form the sulfur cycle. The 'mini-sulfur cycle' is the quick 2-electron transition between sulfide and elemental sulfur (S^o, redox state = 0). Anoxygenic phototrophic (photosynthesis) or colorless sulfur bacteria (chemosynthesis) oxidize sulfide to elemental sulfur and anaerobic respiring bacteria reduce it back to sulfide. Eventually, anoxygenic phototrophic or colorless sulfur bacteria oxidize sulfur further to sulfate. Colorless sulfur bacteria may also oxidize sulfide directly to sulfate. Sulfate can go three ways: (i) it can be assimilated for synthesis of structural cell material, (ii) it can be reduced to sulfide by anaerobic respiring sulfate-reducing bacteria, and (iii) it may be deposited as organic (and inorganic) minerals in sediments. Sulfide is re-mineralized from dead organic matter (recognized by the smell of rotten eggs). During these sulfur transitions several intermediate redox states may be formed. Bacteria may use these intermediates (e.g. S^o, SO_3^{2-}, $S_2O_3^{2-}$) by oxidizing part of it to sulfate and using the electrons generated to reduce the same component to sulfide. This process is also known as inorganic fermentation or disproportionation. (Courtesy Lucas Stal)

2.2.2.4 Phosphorus

Phosphorus is an essential micronutrient in organisms. It is an important component of nucleic acids (DNA and RNA) and therefore central to cell division, transcription and translation processes and protein synthesis. With its energy-rich bond in nucleotides (ATP) it is also central to cellular energy metabolism. Phosphorus occurs in various redox states, but unlike carbon, nitrogen, sulfur and iron, it is not subject to dissimilatory metabolism. Phosphorus occurs in the environment as phosphate. It is taken up as such and occurs in the cell also in the oxidized state, from which it is liberated after cell lysis. Recently, phosphonate has been recognized as a source of phosphorus in the marine environment. There are various enzymes such as phosphatases, nucleotidases and phosphonatases that are involved in the degradation of P-rich organic compounds. Enzymatic hydrolysis of organic phosphorus is an essential step in the biogeochemical phosphorus cycle. Bacterioplankton in low-P environments (e.g., the Sargasso Sea, the Mediterranean Sea) are particularly frugal with P. Phospholipids and nucleic acids appear to be the primary cellular reservoirs of P in the open sea. Recent studies show that plankton have evolved mechanisms to economize on their biochemical P requirements. For example, *Prochlorococcus* and *Synechococcus*, the picocyanobacteria that often dominate low-P environments, primarily synthesize sulfur-containing membrane lipids rather than phospholipids. This switch from P- to S-lipids obviously decreases the cellular demand for P and may be an important adaptation to low-P environments. Phosphate moves quickly through the biosphere but the processes that move it through sediments and the water column are slow, making the phosphorus cycle overall one of the slowest biogeochemical cycles. Phosphates are often trapped in marine sediments as iron or calcium minerals and may be liberated through the activity of anaerobic bacteria.

2.2.2.5 Iron

Iron, one of the most abundant elements on Earth, is a micronutrient essential for virtually all living organisms. It occurs in iron-sulfur cofactors in various enzymes and electron transport components (cytochromes). Oxidized (ferric) iron is virtually insoluble except at very low pH. Reduced (ferrous) iron, in contrast, is soluble but is readily oxidized under aerobic conditions. Hence, bioavailable iron will be present at very low concentrations in the water column of the ocean and is, therefore, suspected to be the limiting nutrient for primary production.

Sahara dust containing large amounts of iron is regularly transported by wind to the Atlantic Ocean and the Mediterranean Sea and has been thought to fertilize these marine environments stimulating primary production and N_2 fixation (see Information Box 4). It is possible that UV light reduces dust ferric iron before it is taken up by the phytoplankton. Alternatively, microorganisms can excrete organic iron chelating compounds called siderophores that bind iron and transport it into the cell.

Information Box 4. Sahara dust and marine microbial diversity

Sandstorms, or dust storms, are usually the result of atmospheric convection currents, which form when warm, lighter air rises and cold, heavier air sinks. Dust from the Sahara Desert can be transported over thousands of kilometres by convection currents, which also cause other meteorological conditions such as thunderstorms (Figure 2.8). Sandstorms are common over the Sahara, which is a major source of mineral dust, transported to the Mediterranean, the tropical Atlantic and the Caribbean. Saharan dust plays an important role in the Mediterranean region because it is the major source of mineral nutrients for phytoplankton and bacteria. However, it is not always beneficial. In the Caribbean, Saharan dust is believed to infect coral reefs with the sea fan disease and some studies suggest dust is linked to health risks, such as increased incidences of paediatric asthma attacks and epidemics of lethal meningitis in the semi-arid sub-Saharan territory known as the Sahel belt, although more research is needed to confirm this.

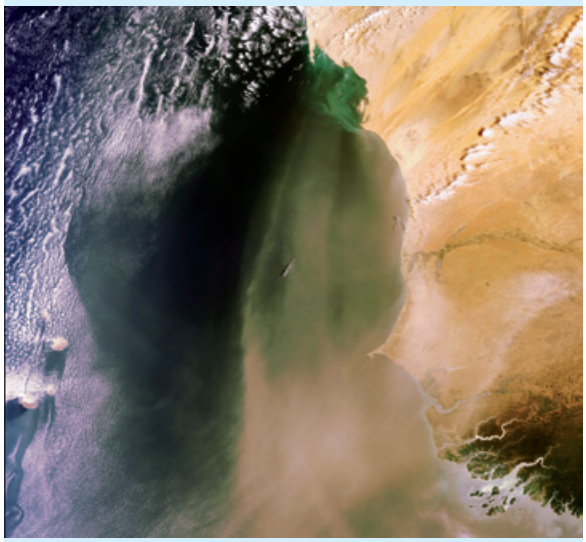


Figure 2.8. ESA Satellite Envisat captures sand and dust from the Sahara Desert blowing across the Atlantic Ocean along the coasts of Mauritania (top), Senegal (middle) and Guinea Bissau (bottom). The cloud-covered Cape Verde islands are visible off the coast of Senegal. (©ESA)

Iron was proposed to be the nutrient limiting factor in the so-called high nutrient low chlorophyll areas, such as the Southern Ocean or the equatorial Pacific. To test this hypothesis a series of large scale iron fertilization experiments were designed to increase marine biological productivity and test if atmospheric carbon dioxide could be sequestered into the deep ocean. Results confirmed that large phytoplankton blooms can be artificially created, at least temporarily.

Oxidized and reduced iron are subject to microbial conversions. Oxidized iron can serve as an electron acceptor for anaerobic respiration and a variety of microorganisms are capable of this reaction. It will be clear that these microbial transformations and dissimilatory iron metabolisms are confined to specific environments, especially sediments, OMZs and certain geothermal springs. Reduced iron is subject both to microbial and chemical oxidation. Iron forms minerals with a variety of other elements, particularly with sulfur and phosphate.

2.2.3 Ecosystem functioning

Although there are ecosystems that function independently from sunlight, some of them in the marine environment, the vast majority of life on Earth (including the ocean) depends on photosynthetic primary pro-

Information Box 5. Iron fertilisation

Iron fertilization, the intentional introduction of iron to the upper ocean to stimulate a phytoplankton bloom, has long been proposed as a potential way to geo-engineer the removal of carbon dioxide from the atmosphere. This stems from the observation that growing plankton takes up carbon dioxide during photosynthesis and iron is often the limiting nutrient for phytoplankton growth in the open ocean. When phytoplankton die, they sinks to the bottom of the ocean locking away some of the carbon they has absorbed from the atmosphere. Fertilization also occurs naturally when upwelling brings nutrient-rich water to the surface or when weather carries dust long distances over the ocean, or iron-rich minerals are carried into the ocean by glaciers, rivers and icebergs.

Since the early 1990s, a number of research groups, scientists and private organizations have been exploring iron fertilization as a means to stimulate phytoplankton growth with the aim of sequestering atmospheric carbon dioxide in the deep ocean. This research has confirmed that large phytoplankton blooms can be created by artificially supplying iron to iron-deficient ocean waters. However, iron fertilization remains highly controversial in terms of the effectiveness of atmospheric CO_2 sequestration and the associated ecological effects. Computer models suggest that iron fertilization of the oceans is much less efficient at decreasing atmospheric carbon dioxide than initially hoped. Opponents of this proposed geo-engineering strategy cite the risk of unintended environmental impacts and argue that in addition, such schemes could distract from efforts to decrease anthropogenic carbon emissions.

Scientists acknowledge that much remains to be learned about the efficiency of iron fertilization in promoting long-term sequestration of carbon dioxide by the oceans, as well as its impact on marine ecosystems. For this reason, a new impetus to iron fertilization research has taken shape under the form of the In situ Iron Studies (ISIS) consortium, an international effort established in February 2011, to assess the efficiency of ocean iron fertilisation (OIF) and determine its potential impacts on marine ecosystems. For more information see also http://noc.ac.uk/news/international-consortium-study-impacts-iron-fertilisation

Figure 2.9. Envisat satellite image showing green swirls of a phytoplankton bloom in the North Sea off the coast of eastern Scotland (©ESA)

duction. In the marine environment virtually all primary production is attributed to microalgae (*Eukarya*) and *Cyanobacteria* (*Bacteria*) (each roughly responsible for half of the global oceanic primary production). In shallow (coastal) waters we may find macroalgae and seagrasses that locally contribute importantly to primary production but globally, they are of minor importance in terms of CO_2 fixation.

Primary production depends on sunlight and is, therefore, restricted to the upper photic zone of the ocean, usually defined as the depth at which sunlight is attenuated to 1% of the surface incident intensity. The depth to which it reaches depends on the absorption characteristics of the water column and may range from 200m in the open ocean to centimetres in some turbid estuaries. The 1% rule does not exclude the possibility of primary production at lower values.

Primary production is the fixation of CO_2 into organic carbon using sunlight as the source of energy and water as the source of electrons, resulting in the evolution of O_2 (oxygenic photosynthesis). The organic carbon and the nutrients (N, P, micro- and trace nutrients) that the primary producers have accumulated form the basis of the oceanic food web. Zooplankton graze on the primary producers, are grazed themselves by other zooplankton and fish and so up to higher and bigger predators. Each of these groups produce fecal pellets which contain organic matter and nutrients that serve as a substrate for *Bacteria* and *Archaea* that decompose it and recycle the nutrients. Living and dead primary producers, fecal

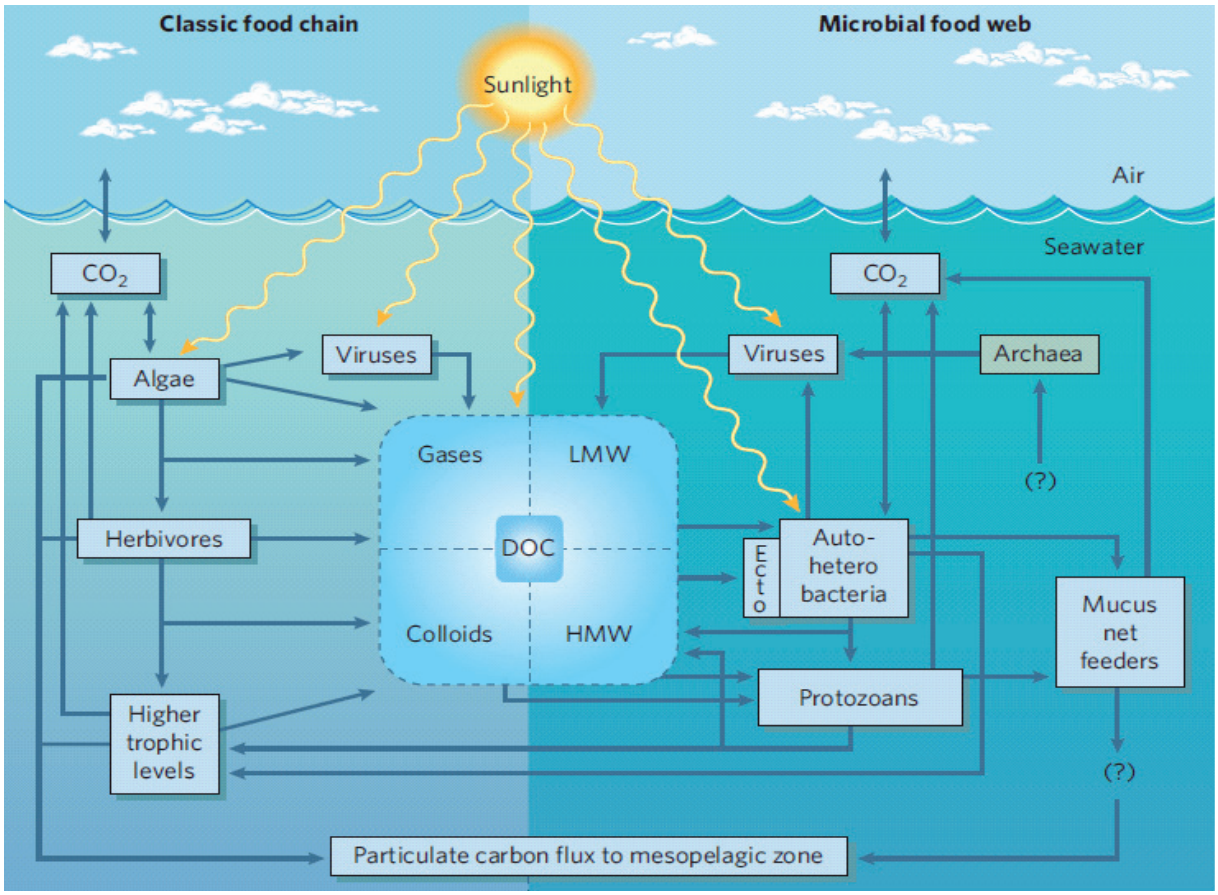


Figure 2.10. Schematic representation of the ocean food web showing, on the left, the classic pathway of carbon and energy flow through photosynthetic *Eukarya*, to herbivores and on to higher trophic levels. Depicted on the right is the microbial food web, which uses energy stored in the non-living, detrital carbon pool to produce microbial biomass that can re-enter the classic pathway of carbon and energy flow. Cell-associated ectoenzymes (Ecto) enable bacteria to use high-molecular-weight (HMW) dissolved organic carbon (DOC) in addition to the more traditional low-molecular-weight (LMW) and gaseous carbon substances. Also shown in the microbial food web are viral particles and *Archaea*. At the present time, there is only rudimentary knowledge of the role of *Archaea* in the oceanic food web. Shown at the bottom of this diagram is the downward flux of particulate carbon (and energy), which is now thought to fuel most subeuphotic zone processes. The classic algae-herbivore grazer pathway (left side) is most important in this regard. (From DeLong and Karl, Nature 437, 2005)

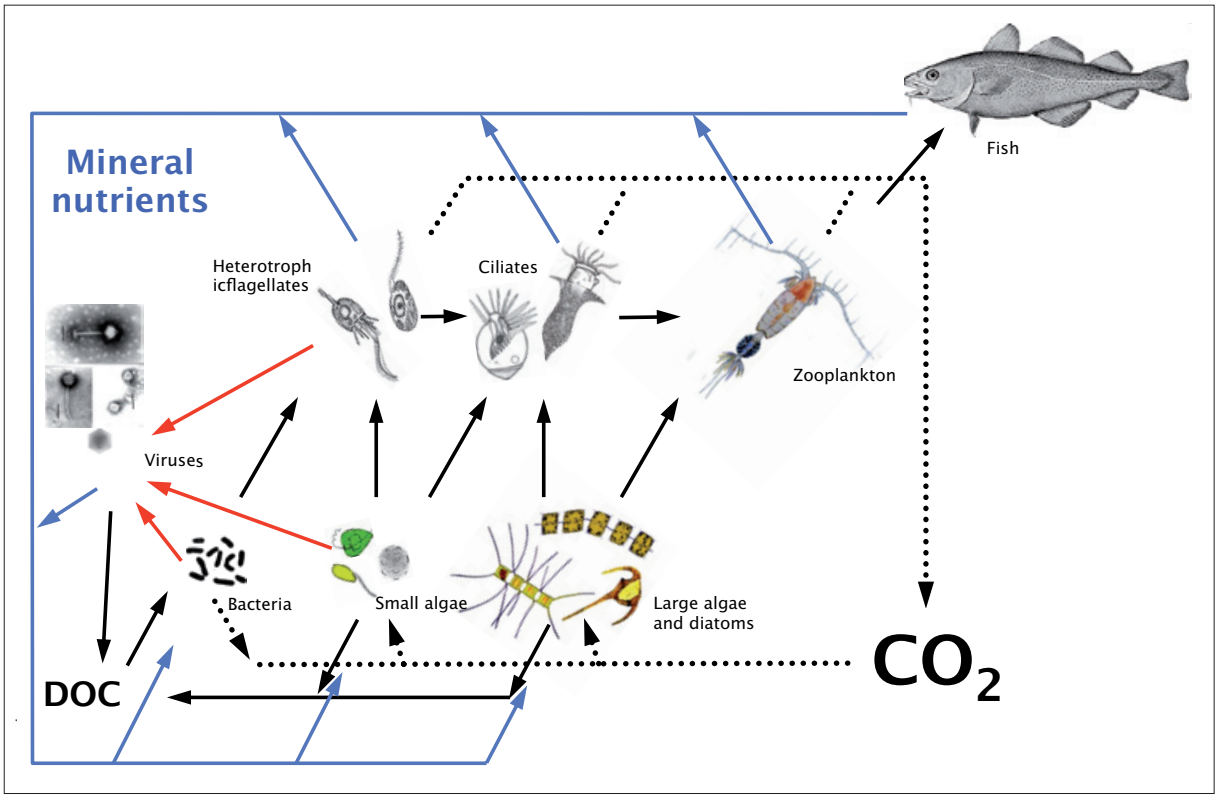


Figure 2.11. Microbial food web and virus-mediated carbon flow (the microbial and viral loops). The red arrows represent virus-mediated pathways. The black arrows represent transport of dissolved organic carbon (DOC) in the microbial food web. The black dotted arrows show the contribution to CO_2 cycling and the blue lines the transport of mineral nutrients in the microbial food web. DOC is the largest biogenic pool of carbon in the ocean. The dynamics of DOC may have an indirect impact on the global carbon cycling contributing to control of atmospheric CO_2. DOC is only accessible to microbes, primarily heterotrophic bacteria. When viruses lyse the host, carbon is transformed from a particulate form (POC) into the dissolved form (DOC). (Modified from Sandaa, 2009)

pellets and aggregates of debris, organic matter and microorganisms sink and provide the food web below the photic zone with substrate. Microorganisms may also escape the food web by direct lysis through viral infection (viral loop) or through exudation/excretion processes through which the organic carbon and nutrients become directly available to the microbial community.

Most of the biomass and biological activity in the ocean is microscopic with microbes playing a central role in Earth's carbon, oxygen, sulfur, and nitrogen cycles. Microbial communities are much more important in the overall marine food web than was assumed until recently. Great numbers of micro-, nano-, and pico-plankton make up the microbial food web which is linked in many ways to the classical food chain with dissolved organic carbon playing an important central role (Figure 2.10).

2.2.3.1 The role of viruses

Viruses are known to be the most abundant and diverse organisms on Earth. Understanding their evolutionary processes is, therefore, of fundamental importance. Viruses contribute to biogeochemical cycles by lysing up to 30-50% of the microbial biomass every day. Carbon and nutrients are released upon viral lysis and, hence, they mediate and enhance biogeochemical cycling in the marine environment (Figure 2.11). In addition, viruses directly affect the abundance and diversity of host cell communities and contribute to microbial gene exchange, which are important for the overall evolution of both the host and the viral community. We are beginning to unravel the significance of viruses and viral-mediated processes in the ocean and considerable effort has been targeted at uncovering viral diversity e.g. through metagenome sequencing. Also, the use of novel techniques in the study of virus-host interactions in culture has uncovered new and unexpected results. Virus-host interactions have revealed that viruses may

also have a tight and mutualistic relationship with their hosts. Indeed, 'host' genes have been uncovered in the viral genome, such as those encoding components of the photosynthetic apparatus, that are likely beneficial to the viral infection process. At present, several genes with hitherto unknown function have been identified that might play a role in the virus-host relationship. This new knowledge about virus-host interactions indicates that viral genes could play important roles because they may affect critical rate-limiting steps of the host metabolism. Nevertheless, our understanding of the effect of viruses on marine community structure and rates of nutrient cycling is far from complete, and their role in the marine ecosystem is still emerging as methodologies evolve.

2.2.3.2 The importance of chemosynthesis in the ocean

There are several sites of chemosynthetic CO_2 fixation in the ocean. Chemosynthetic microorganisms (chemolithoautotrophs) use the oxidation of a reduced compound to generate energy with which they fix CO_2. Although this process has been considered as of minor importance in the water column, Recent reports indicate localized chemosynthesis may be more important than previously believed. For instance, planktonic *Thaumarchaeota*have been shown to oxidize ammonium and fix CO_2 and such organisms may occur in considerable numbers. In oxygen minimum zones (OMZs), cold and hot seeps and black smokers, life is formed through chemosynthetic CO_2 fixation at the expense of the oxidation of sulfide, methane or other reduced compounds (e.g. iron, manganese). These ecosystems are fully independent of light and organic matter which is unique for marine systems.

2.2.3.3 The critical role of the fixation of dinitrogen in the ocean

A critical nutrient for primary production is nitrogen. Although dinitrogen represents by far the largest source of nitrogen on Earth, it is unavailable to virtually all organisms. This is due to the very stable triple bond between the two nitrogen atoms rendering N_2 almost inert. Only some *Bacteria* and a few *Archaea* are capable of reducing N_2 to ammonia and assimilating it. Nitrogenase, the enzyme responsible for N_2, functions only under anaerobic conditions, limiting its activity to anaerobic environments, anaerobic microorganisms or to organisms that provide low O_2 conditions for the enzyme. In the ocean, *Cyanobacteria* are the major N_2-fixing (diazotrophic) organisms and they are responsible for 50% of the global fixation of N_2 (excluding anthropogenic fertilizer production). As oxygen-evolving phototrophs, *Cyanobacteria* have a problem with the incompatibility of N_2 fixation with O_2.

The non-heterocystous, filamentous *Cyanobacteria Trichodesmium* spp. was previously considered the major diazotroph in the ocean. These organisms occur only in the (sub)tropical ocean where the water temperature is well above 25°C. However, recently it was found that unicellular diazotrophic *Cyanobacteria* may be equally important in some oceanic regions. The most abundant belong to the very small (~1 µm) uncultured 'UCYN-A' group which seems to lack the oxygenic photosystem II. These organisms may live in a symbiotic relationship.

The apparent absence of N_2 fixation from the temperate and cold regions of the ocean is not fully explained. It is generally assumed that the marine environment is chronically depleted of combined nitrogen and we would, therefore, expect that N_2 fixation is important here as well.

2.2.3.4 Conversion of bioavailable nitrogen to dinitrogen

The traditional pathway of ammonium oxidation in the marine (but also freshwater and terrestrial ecosystems) nitrogen cycle follows aerobic oxidation in two steps via nitrite to nitrate. Two specific groups of *Bacteria* are responsible for each of these steps. The oxidation of ammonium to nitrite is carried out by ammonium oxidizing (nitritifying) *Bacteria*. However, there is compelling evidence that in the marine environment, certain *Thaumarchaeota* are in fact oxidizing the bulk of ammonia to nitrite. The oxidation of nitrite to nitrate is carried out by nitrite-oxidizing (nitratifying) *Bacteria*. It is not clear how important this second oxidation step is in the marine environment.

Nitrate can follow two routes. It can be used as a nitrogen source by photoautotrophic plankton (known as 'new' or recycled nitrogen) or under anaerobic conditions can be denitrified to N_2 or reduced back to ammonium (DNRA). We can anticipate that the microbial conversions of nitrogenous compounds are quick. While both denitrification, DNRA, or nitrate assimilation go through nitrite as the first intermediate, it might not be unreasonable to conceive that the oxidation of nitrite to nitrate is of limited importance in the marine environment and that nitrite is readily used in these processes before being oxidized to nitrate. It would also explain why ammonium-oxidizing *Thaumarchaeota* are so abundant while this is not so clear in the case of the oxidation of nitrite to nitrate. Apart from being assimilated or used in denitrification and DNRA, nitrite is also a substrate in a relatively recently discovered process, anaerobic oxidation of ammonium (anammox), in which it is used as the oxidant. Anammox oxidizes one molecule of ammonium with one molecule of nitrite, resulting in the formation of one molecule of N_2. It is, therefore, next to denitrification, the second process leading to conver-

sion of 'fixed' and bioavailable nitrogen to the largely inert atmospheric dinitrogen. Anammox bacteria belong to the *Planctomycetes*.

Anammox has been discovered in marine ecosystems, including sediments as well as in the water column in OMZs, but its quantitative importance is under debate. Only one genus, *Scalindua*, has been found in natural marine environments. It seems that in sediments denitrification is more important but in the water column of OMZs, anammox may be the dominant process. In the marine suboxic water column anammox is not particularly abundant and represents <4% of the microbial community. These organisms grow slowly with a doubling time of ~11 days. Anammox is sensitive to oxygen and also to sulfide, explaining its absence in environments dominated by sulfate reduction. Anammox bacteria are autotrophic organisms that fix CO_2 through the acetyl-CoA pathway. This chemosynthetic CO_2 fixation is probably not of quantitative importance for the global C cycle.

2.2.3.5 Summary conclusions: marine microbial research on ecosystem functioning

For decades scientists have invested much time and effort in investigating oceanic biogeochemical processes and the microbes responsible for them. However, new discoveries in the past few years, as a result of the development of new genomic and geochemical technologies, have dramatically changed our understanding of biogeochemical cycling processes at different levels. Examples are the discovery of the processes of anaerobic ammonium oxidation (anammox), and the discovery of ammonia oxidation by *Archaea*, recognized as two new fundamental links in the global marine nitrogen cycle. Thus, there is a need to move from the description of organisms to functional analysis using methods that measure and monitor microbial functions. Moreover, to gain a full understanding of the contemporary and the probable future state of the marine environment, we need to understand the genetic basis of marine microbial biogeochemistry and ocean processes.

2.2.4 Global change

Current global change is associated with a range of physical and chemical factors (rising air and water temperatures, increasing CO_2 uptake by the oceans and seas causing acidification of the marine environment, sea level rise, eutrophication, etc.) and it has been shown that these processes have a profound impact on marine ecosystems, an impact that is predicted to increase (Figure 2.13).

Since the beginning of life, microorganisms have left a notable fingerprint on the geochemistry of Earth and its biogeochemical cycles. Earlier microbes have shaped our atmosphere under which more complex life developed, paving the way for the advent of multicellullar organisms and complex biological communities. While microbial communities will adapt to global change, they might modify ecosystem functioning in a way that we find undesirable.

Climate change is usually referred to as the increase in the average global temperature as the result of an increase of greenhouse gases in the atmosphere (usually emphasizing CO_2) due to anthropogenic activities. While climate changes are a natural phenomenon on Earth, the change we are currently anticipating is unprecedented both in the speed by which it is likely to occur and the expected size of this change.

2.2.4.1 Consequences of a rise in temperature

An increase in global temperature will cause a rise in sea surface temperature (SST). Temperature has a direct effect on the growth rate and metabolic activity of marine microorganisms. It may also select for different microorganisms that are better adapted to higher temperatures or existing microorganisms may disappear when temperatures move outside their normal tolerance range. As an example of the type of processes involved, respiration rates may change with temperature differentially compared to production rates, which could drive the ecosystem towards a low nutrient status (oligotrophy). Similarly, bacterial production and loss rates to grazers might also be affected differentially by temperature, so that an increase in temperature might tighten predator control. Increasing temperatures will also affect the density of the water and consequently stratification and currents. This will, in turn, affect the transport of nutrients. Another effect of the increase of the global temperature is an expected increase in the number and intensity of storms. These events may cause deep mixing of the water column of oceans and seas and may counteract stratification. The sum of these processes is complex and difficult to predict.

2.2.4.2 Consequences of a rise in atmospheric CO_2

The anthropogenic emissions of CO_2 to the atmosphere cause a rise in global temperature (greenhouse effect) with a range of direct and indirect impacts on the marine environment including changes in sea water temperature, sea level, geographic distribution and abundance of species, ecosystem functions, sedimentation processes, etc. In addition, the associated increased uptake of CO_2 has sparked ocean acidification (see below). The effects and impacts of these changes on ocean microbial processes are far from understood. Thus, one of the challenges of the future research will be to understand how climate changes may influence microbial life in the

ocean but also the other way around, i.e. how marine microbes influence climate.

The increase in atmospheric CO_2 concentrations may affect marine microorganisms in various ways. In the first place, it might stimulate primary producers, although an increase in CO_2 is the result of a complex cascade of processes. For instance, an increase in primary production would only be expected when CO_2 is indeed the limiting factor. This is unlikely to be the case for many primary producers and for many prevailing environmental conditions. Indeed, even if primary production were enhanced, the resulting phytoplankton blooms, even if not toxic, often result in poor water quality. Decomposition of the increased organic matter would lead to oxygen depletion and to anaerobic processes that in the marine environment would lead to the formation of toxic sulfide. The cascade of such processes may in fact move an ecosystem state into another, unwanted, stable state.

The oceans are a sink of CO_2 by a process that is known as the biological pump. The CO_2 that dissolves in the water is used by phytoplankton to be converted into biomass. Dead biomass sinks and is decomposed by bacteria which produces CO_2 and the seawater at high depths (with higher pressure) contains considerable higher concentrations than surface water which is in equilibrium with the atmosphere. Another (small) part of the organic matter is eventually buried in the seafloor and recycles on geological time scales. Calcium carbonate skeletons of coccolithophoric algae (coccoliths) rain down to the seafloor and also this inorganic carbon is buried. The big unknown is how much CO_2 can be accommodated by the oceans and what would happen if the CO2-rich water masses come to the surface.

2.2.4.3 Acidification

The other effect of increasing CO_2 concentration is known as ocean acidification. CO_2 is an acid and its dissolution in seawater will decrease the pH. The effects of small changes in pH on the growth and activity of marine microorganisms are probably minor, except for calcification which is negatively affected so that the dissolution of calcium carbonate (such as that forming coral reefs) may occur. Calcification is essential for shellfish and an effect of acidification on shell formation has been demonstrated. It is not clear whether coccolithophorid algae are affected by acidification and although deformation of the coccolith platelets has been demonstrated, it is uncertain whether they are important for the organism except for generating CO_2 (which would be unnecessary in case of high CO_2). More problematic could be the massive dissolution of calcium carbonate, which would increase the dissolved CO_2 and further decrease pH generating a possible run-away effect.

2.2.4.4 Emission of nitrogenous greenhouse gases

Microbial activity also affects atmospheric concentrations of the greenhouse gas nitrous oxide (N_2O). N_2O has been identified as the dominant ozone-depleting compound and it is projected that it will continue to be throughout the 21st century. Thus, understanding the processes controlling emissions of N_2O from coastal and marine systems is important for evaluating climate change scenarios.

The source of N_2O in the ocean is most likely incomplete denitrification, although nitrification may also lead to the formation of N_2O. Denitrification occurs under anaerobic conditions, in aggregates (marine snow), OMZs and sediments.

While we would anticipate that nitrate is in low supply relative to the organic matter that is oxidized, it is unclear why the nitrogen is not oxidized all the way to N_2. It is possible that other sources of N_2O exist. For instance, it has been suggested that dinitrogen-fixing organisms may produce it. The only sink of N_2O is the last step in denitrification where it is reduced to N_2. However, if blocked, for example due to the presence of Dimethyl Sulfide (DMS) or other sulfur components, N_2O accumulates. Once in the water column it is most likely emitted to the atmosphere. Increasing the load of nitrogen by dry and wet deposition, run-off and riverine discharge, as well as by the fixation of dinitrogen, may lead to processes that increase N_2O emissions.

2.2.4.5 Dimethyl Sulfide (DMS)

Oceans emit considerable quantities of dimethylsulfide (DMS), which is produced from the degradation of dimethylsulfoniopropionate (DMSP) by *Bacteria* and microbial *Eukarya*. DMSP is produced by some microalgae and serves as an osmoprotectant, allowing them to grow at high salt concentrations (in seawater). The emission of DMS into the atmosphere is thought to act as a climate feedback regulation. The so-called CLAW hypothesis states that DMS in the atmosphere is converted to an aerosol form of sulfate (non-sea salt sulfate, NNS) and that these aerosols act as cloud condensation nuclei (CCN). Hence, the more DMS is emitted, the more cloud albedo is formed, resulting in a greater reflection of sunlight. This would subsequently decrease the growth of microalgae and the production of DMSP and consequently less DMS would be emitted. It is, however, likely that this climate feedback is far more complex. DMS may also be the transporter of sulfur from the oceans to the continents were it precipitates as sulfuric acid with rain. This may enhance rock weathering, securing the transport of nutrients from the continents to the sea, nutrients which support microalgal growth and primary

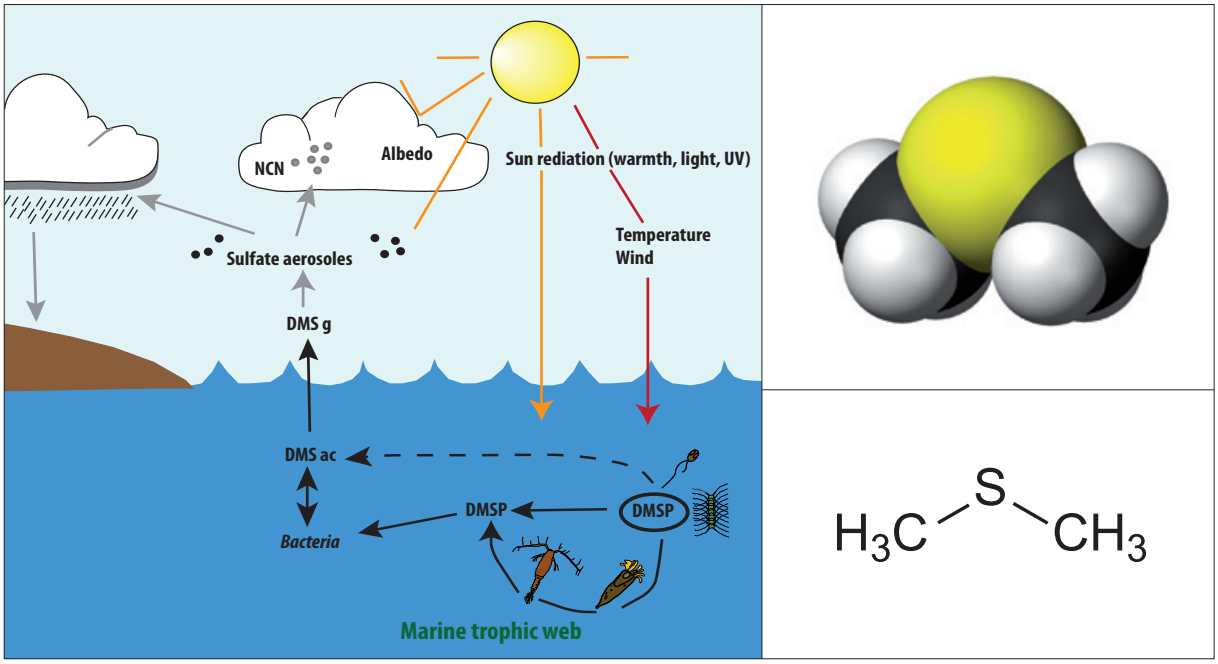


Figure 2.12. Dimethyl sulfide (DMS) or methylthiomethane is an organosulfur compound with the formula $(CH_3)_2S$ (right) which provides the characteristic smell of seafoods and when cooking certain vegetables, notably maize, cabbage and beetroot. The production of dimethylsulfide (DMS) in the oceans is driven by microorganisms (left; courtesy JM Gasol and R Simó). The emission of DMS into the atmosphere is thought to act as a climate feedback regulation.

production. These are all interesting hypotheses that emphasize the complexity of the biogeochemical cycles on Earth.

2.2.4.6 Methane

Methane is a very potent greenhouse gas. There are basically two sources of methane; one is geothermal (natural gas, volcanic, hot springs, cold seeps, methane gas hydrates) and the other is biological. Methanogenic *Archaea* are anaerobic organisms that convert CO_2 or simple organic compounds (e.g. acetate) to methane. In terrestrial environments these organisms are particularly known from the rhizosphere of rice paddy fields and from the intestines of a range of organisms such as cattle and termites. Anthropogenic activities have increased this source of methane considerably. Methanogenic bacteria are also known from swamps and wetlands. However, in the marine environment methanogenic bacteria compete with sulfate reducing bacteria which have an advantage because of the very high concentration of sulfate in seawater (28 mM). Therefore, methanogenic bacteria may play a less prominent role in the marine environment, even though there are non-competitive substrates (which are not used by sulfate reducing bacteria) and molecular surveys indicate their abundant presence. Moreover, methanotrophic bacteria are expected to utilize any methane that is formed. In the water column aerobic methanotrophic bacteria would consume methane, while under anaerobic conditions, consortia of methanogenic bacteria and sulfate reducing bacteria oxidize methane (in terrestrial anaerobic environments methane is oxidized by oxygen which the organism generates from nitrite). Seafloor gas hydrates are a possible source of methane also. Such massive amounts of methane will travel through the water column and be released into the atmosphere without being consumed. These are phenomena independent of human activities.

In conclusion, there are still many unknowns regarding sources and fates of methane in the sea and the role of marine microorganisms in methane regulation. Much of our knowledge is based on old paradigms which may need revision. The possible emission of methane from coastal swamps and wetlands needs particular attention.

2.2.4.7 Freshwater inflow

Increasing global temperature is causing melting of the polar icecaps resulting in sea level rise. Apart from the problems this will cause in countries bordering the sea, the melting of ice may have a range of other effects. The sudden release of massive volumes of freshwater

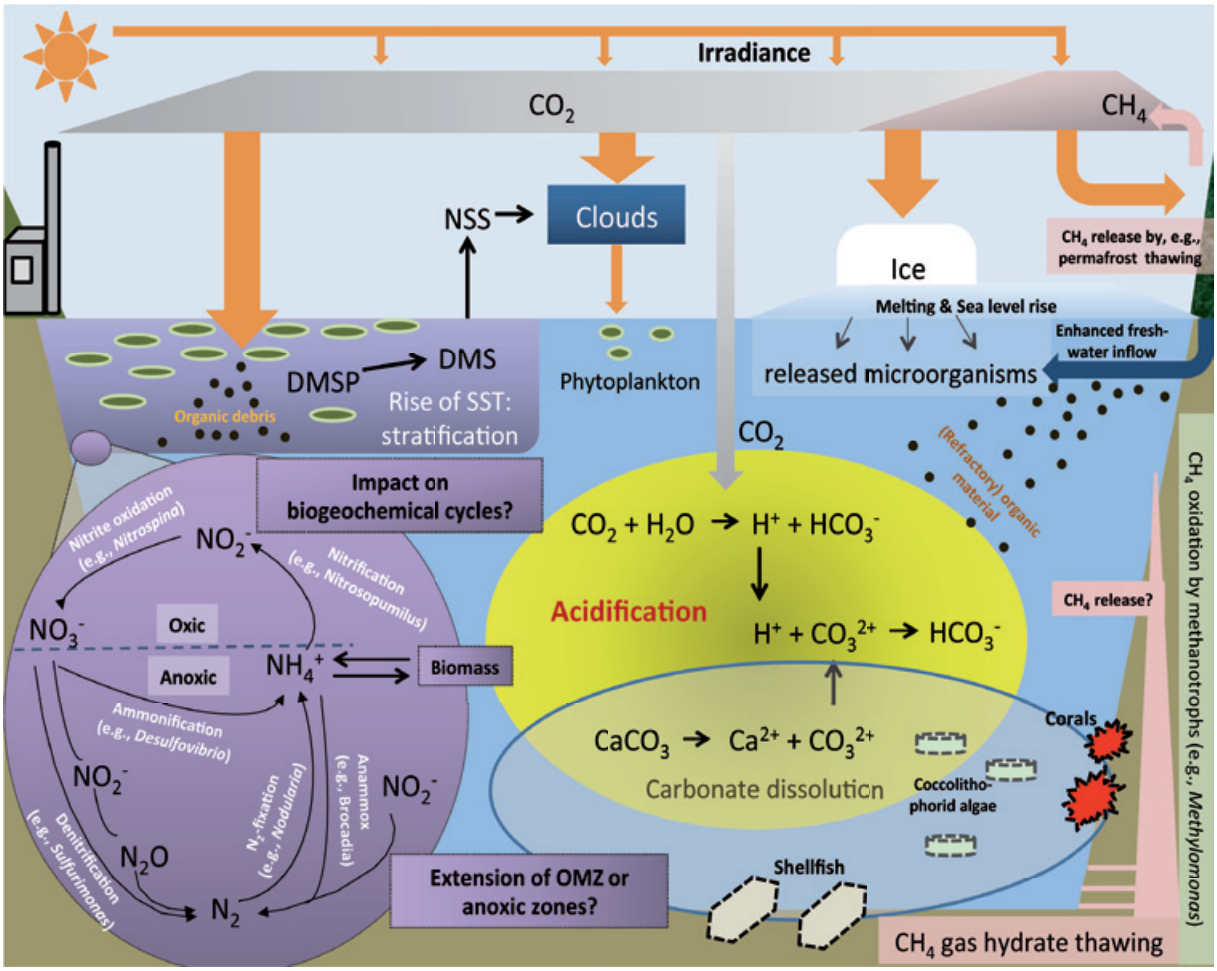


Figure 2.13. Scheme illustrating the potential global change effects on the ocean. Globally increasing temperatures and concentrations of greenhouse gases such as CO_2 and CH_4 will have significant impacts on microbial activities and communities, which may affect biogeochemical cycles. (Developed by Matthias Labrenz)

will cause salinity changes which may affect water currents and stratification. Obviously, salinity changes also affect microorganisms and may cause dramatic shifts in the composition of the microbial community. In addition, with the melting of the ice, trapped organic matter, contaminants and microorganisms are released and these could also cause shifts in local community composition.

2.2.4.8 Summary conclusions: marine microorganisms and global change

Gas exchange between the ocean surface and atmosphere involves air (mostly oxygen and dinitrogen), and the trace gases CO_2 (carbon dioxide), N_2O (nitrous oxide or laughing gas), CH_4 (methane), and dimethylsulfide (DMS). Some of these gases are greenhouse gases and considered to be responsible for global warming. In addition, the increasing dissolution of CO_2 leads to acidification of seawater, which could cause the dissolution of carbonates. DMS may trigger cloud albedo formation and cause acid rain. These gas exchanges are the result of biogeochemical processes which depend on the composition and activity of the marine microbial community, which in turn respondss to and cause global climate change. Unfortunately, our understanding of these complex processes and interactions is limited. We are just beginning to collect the necessary genetic data, discover the relevant biochemical processes, and get a grip on the untold diversity of marine microorganisms and their interactions. This will be the basis of predictive models which will help us to better manage the ocean as one of our most precious and vulnerable resources.

2.2.5 Local anthropogenic forcing

Aside from global changes, a number of human activities exert an important influence on marine microbial life. Some examples are habitat modification, freshwater runoff and river discharges, oil and other chemical spills. Occasionally, these lead to unwelcome consequences such as eutrophication and the occurrence of harmful algal blooms (HABs). Humans change the coastline and may also modify the characteristics of the substrates thus affecting the role of microbes in the environment. A typical case is the building of harbours which decrease the seawater residence time. Under such conditions, when turbulence is prevented and temperature tends to increase, the development of HABs is favoured. These HABs may comprise species which are a nuisance simply because they affect the colour and smell of the water and are not appreciated by swimmers and tourists or, more seriously, may pose a direct health hazard when they produce neuro- or enterotoxins. While in some cases a direct link between human intervention and ecosystem responses can be demonstrated, more often such links remain elusive or indirect. This problem is well known for the Baltic Sea (*Cyanobacteria*), but HABs likewise affect the coastlines of the Atlantic Ocean and Mediterranean Sea (*Phaeocystis*, dinoflagellates), which in recent years have been subject to increased urbanization (see Information Box 6 dealing with cyanobacterial blooms in the Baltic Sea).

The spread or introduction of non-native species across biogeographic ranges is a relevant case of biotic disturbance. The rapid proliferation of non-native macroalgal species may derive from their fast growth and dispersal, high competition with native organisms, and/or the status of the invaded native ecosystems, altered in many cases by human activities. The transmission of non-native species may drive changes in marine biodiversity and the functioning of marine ecosystems, which may decrease ecosystem resistance and resilience to disturbance. HABs are known to be particularly invasive, but very little is known about particular *Bacteria* or *Archaea* which are thought to have a universal distribution.

Increased or decreased freshwater runoff poses an ecosystem threat as coastal water circulation patterns might be affected and terrestrial runoff may be rich in inorganic nutrients, refractory organic matter, or both. This can affect the heterotrophic/autotrophic balance of the ecosystem, which is mostly in place by virtue of

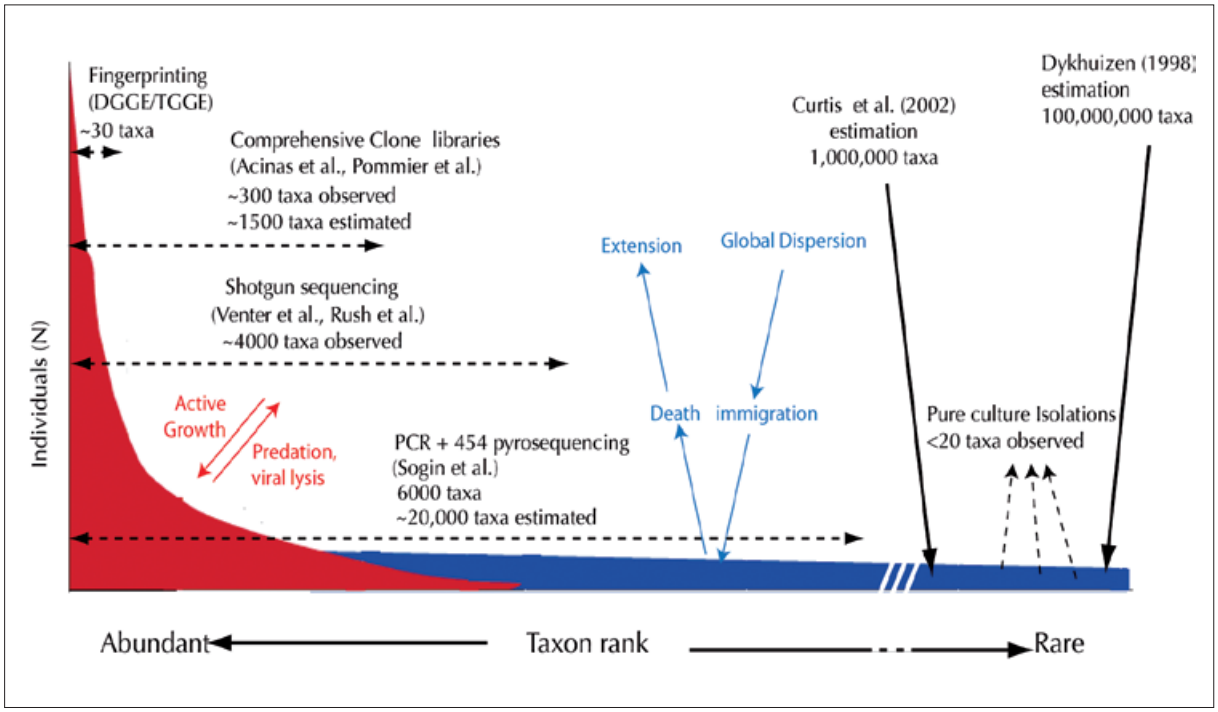


Figure 2.14. A conceptual model of how microbial diversity in the ocean is expressed, i.e. a few organisms are relatively abundant and participate in ecosystem functioning while the majority of microbes await more optimal growing conditions, as shown above in the large 'tail of biodiversity'. The image also indicates the level of diversity that can be assessed by each current methodology (Adaptation by Thomas Pommier, CNRS, of an original figure by Pedrós-Alió, 2006)

microbial action. Similarly, dredging or habitat alteration in the sediments might alter the functioning of sediment microbes by modifying the concentration of terminal electron acceptors (oxygen, nitrate, sulfate, ferric iron) and thus modifying the dominant end product reactions.

The accidental release of man-made organic or inorganic substances which are allochthonous to the marine environment may produce interesting effects. Oil, for example, is known to be degradable by bacteria subject to the presence of sufficient oxygen (for hydrocarbon degradation) and inorganic nutrients. Specialized bacteria are known to reside in the "rare biosphere" of environments subjected occasionally to oil pollution (such as harbours, rivers, marinas, but also including the rhizosphere) and to grow and degrade most of the oil components. Other chemicals which can be present in urban waters can be degraded but a problem may arise in the case of discharges of antibiotics in waste water because it may lead to the development of resistances which could eventually be transferred to human pathogens.

2.3 Societal impacts and benefits

Marine microorganisms influence our lives more than we can imagine. First of all, as illustrated in Section 2.2.3 of this chapter on ecosystem functioning, marine microorganisms provide critical ecosystem services in terms of regulating marine food webs, producing oxygen (about 50% of all the oxygen on Earth), biogeochemical cycling, assuring ecosystem integrity, and CO_2 uptake and buffering (carbonate system). Secondly, marine microorganisms also provide important products and processes which are of use to our societies. At the same time, some marine microorganisms are responsible for potentially negative impacts such as the spread of human and fish diseases and the contamination of fisheries and aquaculture products with toxic substances (e.g. resulting from HABs). This section will describe some of the most important societal benefits which we may obtain from marine microbial research in the future.

Information Box 6. Cyanobacterial blooms in the Baltic Sea

Eutrophication is one of the fundamental results of human activity in the Baltic Sea. High nutrient loads, predominantly from intensive agriculture or municipal sewage, cause the formation of extensive blooms of primary producers in spring and summer. The spring bloom is dominated by diatoms and dinoflagellates and leads to a decrease of nitrate and phosphorus concentrations in the water. The summer bloom is dominated by the cyanobacteria, *Nodularia spumigena*, *Aphanizomenon flos-aquae*, *Anabaena* sp., and *Synechococcus* spp., of which the first three are often characterized by their ability to fix dinitrogen. Dinitrogen-fixing primary producers become predominant in the summer bloom when nitrogen availability is limited but phosphorus is still available. The general availability of phosphorus in the water column is a specific feature of the central Baltic Sea and is connected to the anoxic nature of the bottom waters of this area as anoxia enhances the release of phosphorus from the sediment. The phosphorus stock in the water column thus increases and feeds surface primary production. As demonstrated by satellite images (Figure 2.15), cyanobacterial blooms may cover wide areas of the whole Baltic Sea and since some *Cyanobacteria* are potential cyanotoxin producers, the blooms can be hazardous for higher life forms, which in itself reveals the potential socio-economic impact on Baltic riparian countries.

Figure 2.15. Satellite image, acquired on 31 July 2008, by Envisat's Medium Resolution Imaging Spectrometer (MERIS), captures a cyanobacterial bloom in the Baltic Sea (provided by H. Siegel, IOW). (©ESA)

2.3.1 The relationship between biodiversity and environmental stability

A fundamental ecological tenet is that of the relationship between biodiversity and environmental stability. The idea behind such a tenet is that an ecosystem that has a large number of potential solutions to problems (i.e. a large number of species which can act in slightly different ways) has a higher probability of surviving environmental stress (e.g. the introduction of an allochthonous toxic chemical, or a large increase of a nutrient) than an environment which has fewer available choices. It is always possible that a particular microorganism will grow and replace a dominant organism which is affected in such a way that the overall community performance and ecosystem functioning remains mostly unaffected. This is the so-called "insurance-hypothesis" and it arises from initial studies on this issue performed in temperate grassland communities that demonstrated that plant diversity affects ecosystem functioning such as above-ground production, and ecosystem properties such as resistance and resilience.

While we lack direct experimental evidence of this phenomenon, some studies have shown that where greater numbers of microorganisms exist in an environment, more organic C is degraded and the environment fluctuates less. This suggests that maintenance of biodiversity, including microbial biodiversity, is essential for ecosystem functioning and, by extension, the well-being of human societies. Research that supports this concept is limited by the available methodologies that, as discussed elsewhere in this document, reach only a small fraction of the true microbial biodiversity present in the marine ecosystem.

2.3.2 Biodiscovery and bioactive compounds

The marine environment is emerging as a 'gold mine' for novel bioactive compounds with a staggering 1011 new compounds reported for 2009. Marine-derived natural products present an enormous range of novel chemical structures and provide an interesting and challenging blueprint for creating new entities via synthetic chemistry. Marine invertebrates and plants, in particular, represent an environment rich in microorganisms that produce compounds with bioactive properties including antibacterial, antifungal, antiviral, anticancer, antifouling and antibiofilm activities. However, only 1% of these microorganisms can be isolated using traditional culturing techniques, which has been a major bottleneck when mining the marine environment for novel bioactive molecules.

2.3.2.1 Antimicrobial and antifungal

The emergence of multidrug resistant bacteria and fungi, the latter including certain *Aspergillus fumigatus* and *Candida albicans* strains, drives the continuous search for novel antibacterial and antifungal agents. Most natural antibiotics used today originate from soil actinomycetes. However, since the rate of discovery of novel antibiotics of terrestrial origin is declining, other ecological niches, including the marine environment, are being exploited in the search for new antibiotics. Classes of compounds with antibacterial and/or antifungal activity which have been isolated from the marine environment include peptides, sterols, terpenes, alkaloids, and polyketides.

Three classes of antibiotic resistant bacterial pathogens are emerging as major threats to public health: (i) methicillin-resistant *Staphylococcus aureus* (MRSA), (ii) multidrug resistant Gram negative bacteria, including *Escherichia coli* and *Pseudomonas aeruginosa*, and (iii) multidrug resistant *Mycobacterium tuberculosis*. Numerous compounds which could potentially combat these classes of pathogen have been isolated from the marine environment. These include structurally novel compounds such as marinopyrrole A and abyssomicin C with activity against MRSA, the alkaloid cyclostellettamine F with activity against *P. aeruginosa*, and trichoderins, novel aminolipopeptides with anti-mycobacterial activity. Compared with infections caused by drug resistant bacteria, infections caused by resistant fungal pathogens occur relatively infrequently. However, *Candida* species are a common cause of hospital-acquired bloodstream infection and kill 40% of those patients, whereas disseminated *Aspergillus* infections can kill up to 80% of affected patients. Compounds of marine origin with activity against these fungal pathogens include the cyclic depsipeptide kahalalide F and the alkaloid araguspongin C.

2.3.2.2 Antiviral

Viral diseases, such as HIV and influenza A subtype H1N1, are a major threat to human health. Since viruses can rapidly evolve and develop resistance to currently used antiviral agents, discovering new antiviral drugs is of paramount importance. Many classes of antiviral compounds have been isolated from the marine environment, including nucleosides, terpenes, cyclic depsipeptides, alkaloids, macrolides, and polysaccharides. The first commercial antiviral drug, Ara-A, was synthesised based on the structure of the nucleosides spongothymidine and spongouridine which were isolated from marine sponges. Although Ara-A and other synthetic nucleosides have been used to treat herpes simplex virus (HSV) and HIV, few marine antiviral compounds have entered preclinical trials. Examples of such

compounds include the anti-HIV avarol, isolated from the marine sponge *Disidea avara*, and Cyanovirin-N, isolated from the cyanobacterium *Nostoc ellipsosporum*.

2.3.2.3 Anticancer

Molecules with anticancer properties comprise the majority of bioactive molecules derived from marine sources. However, relatively few compounds enter preclinical and clinical trials and only a small group stems from microbes. Although chemically synthesised, ara-C (cytarabine) was designed based on the nucleosides from spongothymidine and spongouridine, originally isolated from the marine sponge, *Tethya crypta*, and is in clinical use for more than 40 years. Today, Yondelis®, a potent anticancer drug developed by Pharmamar from compounds produced by the tunicate *Ecteinascidia turbinata*, is probably the most successful example of how marine natural products can lead to anticancer treatments.

2.3.2.4 Antifouling and antibiofilm properties

Biofouling, the undesirable accumulation of microorganisms, plants, algae, and/or animals on wetted structures, is of great concern in a wide range of applications, ranging from food packaging/storage, water purification systems, marine and industrial equipment, to medical devices. The two main strategies that are used to combat biofouling are to either prevent initial attachment or to degrade fouling biofilms. Several coatings are designed to prevent this initial attachment. However, some applications of the antifouling coating require certain requirements/restrictions. For example, many countries have now imposed a ban on the most effective antifouling coating (organotins) available for marine applications, urging instead the use of non-toxic novel biofouling compounds. In the health care sector, antifouling coatings for medical applications require compounds that are bactericidal and non-toxic to the human body. Marine organisms, in particular seaweeds and marine invertebrates, have proven to be a successful source of antifouling compounds. However, these marine compounds are difficult to obtain in large quantities. To overcome this problem, marine microorganisms are being explored to identify novel biofouling molecules which can be produced in larger quantities. Although the mechanism is unknown, several fatty acids (e.g. 1-hydroxymyristic acid, 9-Z-oleic acid and 12-methylmyristic acid) produced by marine microorganisms have antifouling properties. The most promising evidence came from an experiment in which a coating consisting of 10% fatty acids prevented attachment of micro- and macro-fouling organisms on a panel which had been immersed in the ocean for 1.5 years. Similar results were obtained from a coating containing synthesised alkyl butenolide, after alkylated butenolides isolated from a deep sea

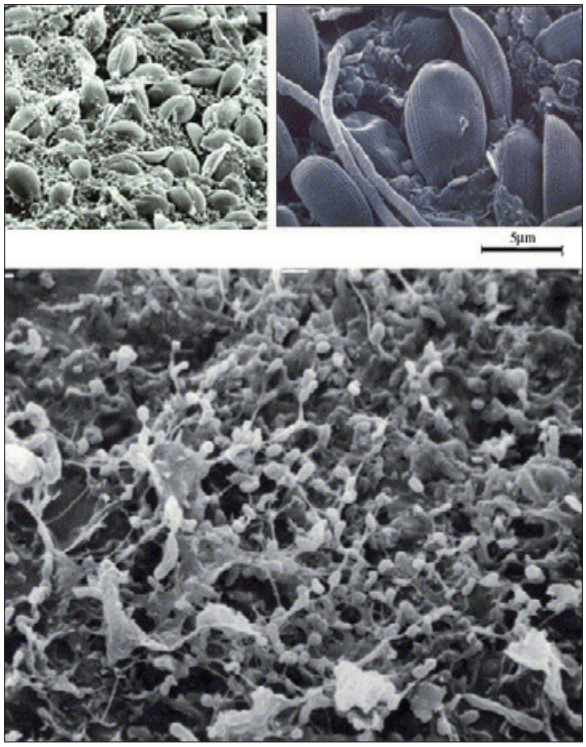


Figure 2.16. SEM picture of a marine biofilm growth on top of a copper-based antifouling paint. The two biofilms shown in the top pictures contain mainly *Amphora* sp. diatom cells settled onto a TBT-SPC paint applied on a ship hull. A thick bacterial biofilm grown onto a vinyl paint containing Cu_2O and TPTF within 4 weeks of exposure to sea water (14,000×, bottom). Pictures courtesy by Dr. Maureen Callow (top left), Hempel A/S (top right) and Dr. Mike Dempsey (bottom).

Streptomyces species were found to exhibit antifouling properties. Several other compounds show promise including pyolipic acid, phenazine-1-carboxylic acid and 2-alkylquinol-4-ones and proteases, but their further potential is yet to be examined.

Bacteria possess a cell-to-cell communication system termed quorum sensing. This communication system is dependent on the production of small molecules, which when sensed in high concentrations lead to coordinated behaviour by regulating several physiological processes including bioluminescence, motility, antibiotic resistance, virulence factor production and biofilm formation. Biofilms (aggregates of microorganisms where cells adhere to each other or to a surface) are of major concern for treating bacterial infections, since bacteria present in biofilms have increased antibiotic resistance profiles compared to their sessile counterparts. Since quorum sensing and biofilm formation are closely linked, it is not surprising that quorum sensing inhibitors often also exhibit antifouling/antibiofilm properties.

Recently, extracts from several microbial isolates from a marine habitat were shown to contain antiquorum sensing and antibiofilm activity against *Pseudomonas aeruginosa*, the primary cause of morbidity and mortality among cystic fibrosis patients. As such, quorum sensing and biofilm inhibitors could provide novel treatment options for treating bacterial infections, and these are urgently needed due to the emergence of multidrug resistant microorganisms.

2.3.2.5 Enzymes

Due to their immense genetic and biochemical diversity, marine microorganisms are of interest as a promising new source of enzymes with unique properties, including salt tolerance, hyperthermostability, barophilicity, and cold adaptation. Proteases represent an important class of industrial enzymes, with applications in the detergent, leather and pharmaceutical industries and many have been isolated from marine microorganisms. Other enzymes that have been isolated from marine microorganisms include lipases (involved in the breakdown of fats and oil), and a range of polysaccharide-degrading enzymes, such as chitinases, alginate lyases, agarases, carragenases, amilases, cellulases and lignocellulases. These latter enzymes are of particular interest in the production of bioethanol, since the development of alternative energy sources is an urgent global priority. Bioethanol derived from crops such as corn are the most developed form of biofuel, but there are also initiatives that aim at the direct production of ethanol by (marine) *Cyanobacteria*. Metabolically enhanced *Cyanobacteria* convert the fixed CO_2 directly into ethanol. Concerns with regard to world food shortages may restrict the long-term viability of plant-biomass derived biofuels. However, biofuel derived from marine microalgae is a promising and clean alternative. Enzymes specific to the breakdown of the complex polysaccharides present in biomass may be isolated from the marine environment. Marine enzymes also have potential applications in the bioremediation of polluted waters.

2.3.2.6 Future prospects: potential of metagenomic approaches

While culture-dependent techniques have been highly successful in the identification of novel bioactive compounds from marine microorganisms, the advent of deep sequencing technologies such as pyrosequencing has revealed a huge diversity of microorganisms in various niches in the marine environment, the majority of which cannot be accessed by traditional culturing techniques. Metagenomics involves the cloning of total community DNA, including DNA from uncultured microorganisms, and subsequent sequence-based or function-based analyses. Sequence-based metagenomic approaches rely on the comparison of cloned DNA to known sequences from databases, while function-based metagenomic approaches involve screening library clones for a phenotype, such as antibacterial or anti-quorum sensing activity. These metagenomic approaches allow access to natural products produced by microorganisms which cannot be cultured, effectively increasing the mining potential of the marine environment.

2.3.3 Bioremediation

Organic contaminants can undergo biodegradation as a result of the activity of microorganisms resulting in less toxic, less mobile and/or less bioavailable products. This is the basis of bioremediation: the use of biological agents, namely microorganisms, to remediate organic contaminants in the environment. Bioremediation is considered an effective, low cost, preferred clean-up option for moderately contaminated areas. Nevertheless, further research is needed to find suitable combinations of microorganisms and environmental characteristics to improve bioremediation processes. In fact, for a remediation process to be effective, the overall rate of degradation needs to be accelerated above current microbial processes. Accelerating the biodegradation of organic contaminants is thus a major challenge to improve the performance and acceptance of cost-saving bioremediation techniques.

A systematic approach to select suitable microbial species for a particular remediation application is crucial to maximize the dissipation of the contaminant in the environment. Hence, the successful application of bioremediation techniques is dependent on the identification and isolation of appropriate microbial strains and their subsequent survival and activity, once released into the target habitat. Strain selection is founded on the principle that certain microorganisms are better suited for particular (catabolic) tasks and environments than others. Increasing our knowledge of the relevant microorganisms and genes will thus be critical to improve biotechnological applications.

2.3.4 Building with nature

Engineers recently begun to appreciate the potential of microorganisms for engineering purposes. Erosion and stabilization of coastal sediments such as intertidal mudflats and sandy beaches may depend, to a considerable extent, on the microorganisms that live in these sediments. For example cyanobacteria form microbial mats in the intertidal areas of sandy beaches. These microbial mats are rigid structures that trap and bind sediment particles and render stability to the sediment, increasing its erosion threshold. Also, as a result of the trapping and binding of sediment, microbial mats tend

to increase the profile of the tidal sediment and give rise to the formation of a salt marsh and eventually new land. Mudflats are characterized by biofilms of benthic diatoms. These photosynthetic microalgae exude copious amounts of extracellular polymeric substances (EPS) which act as a glue for the mud particles, likewise increasing the erosion threshold. Thus, microbial communities can have considerable effects on coastal morphodynamics and enrich and improve the sediment. This knowledge can be used for coastal protection as well as for soil improvement in terrestrial and agricultural environments. Another interesting application of microorganisms is found in so-called self-healing concrete. This concrete is mixed with microorganisms and with a substrate that contains calcium. When the concrete shows cracks, water will enter. Normally this will cause further deterioration of the concrete and oxidation of the reinforcing bars. However, in this case the microorganisms become active, degrade their substrate and precipitate calcite which fills up the cracks. It is obvious that the potential of microorganisms as micro-engineers is far from being fully explored and awaits our imagination and discovery.

2.3.5 Marine food from fisheries and aquaculture

The oceans and seas are an important source of food for a growing global population. Projections from the FAO show an increase in the demand for seafood products worldwide and in Europe towards 2030. The average per capita consumption of 28 European countries included in an FAO foresight analysis indicates an evolution from 22 kg/person/year in 1998 to 24 kg/person/year in 2030. There is general agreement, however, that world fisheries have reached their maximum sustainable yield and many commercial stocks might collapse by 2030 if current trends continue. To supply the growing demand for healthy food from the sea, aquaculture production has increased significantly over the last decades and will need to become even more important in providing sustainable sources of marine food in the future. Marine microorganisms dominate the marine food web and are the basis of the production of the food for all life in the ocean. Not surprisingly, marine microorganisms are of great importance to the fisheries and aquaculture sectors and a better knowledge of their role will be critical to secure the necessary food supply from marine living resources in the coming decades.

Fisheries depend on the adequate functioning of microbial communities. Yet, despite their central role in marine food webs, we still know very little about marine microbial community functioning and how alterations, e.g. as a result of pollution or climate change, might affect commercial fish stocks. This needs to be recognized and addressed.

To meet the challenge of supplying growing seafood markets, aquaculture will need to become more efficient and cost-effective, whilst simultaneously decreasing its environmental impact. Microbiological research has already contributed to improving aquaculture production systems. For example, microbial bioremediation, particularly in land-based mariculture, and microbial control of intensive production systems have improved containment and environmental compatibility. Another critical issue for the commercial aquaculture sector which has been a key area for applied research relates to diseases in cultured fish and shellfish stocks. Usually, bacterial pathogens are treated with antibacterial agents, which can eventually lead to antibacterial resistance with a subsequent and direct impact on human health. A better understanding of host-microbe interactions within confined marine ecosystems may provide us with completely new tools to understand disease development in general and how to fight new diseases in a more ecological way. Disease management through the preventive use of probiotics and immunostimulants such as yeasts is another approach which shows interesting results. Research in this area has mainly concentrated on pathogen exclusion, competition or microbiotamanipulation (see also Marine Board Position Paper 15 on Marine Biotechnology). Microbial control could result in an improved disease management, as well as having positive effects on the environment.

New approaches and techniques such as Integrated Multi-Trophic Aquaculture (IMTA) and Recirculating Aquaculture Systems (RAS) are also becoming increasingly important in the search for more environmentally friendly and efficient production systems. In RAS, for example, seafood production is combined with water purification to maintain a healthy culture environment, often in a closed system. RAS are advanced and complex aquaculture systems that rely on biological processes which are primarily microbial and can, therefore, benefit from advances in marine microbial ecology. In addition, RAS open the possibility to integrate microalgal systems into recirculating aquaculture systems and further downstream in the management of fish processing outflows. Unfortunately, the ecology of microbial communities in RAS and its interaction with the microbiota in the food and gut of cultured organisms is still poorly understood. In addition, the microbiota present during larval development is highly variable and is thought to influence larval viability and health.

Adequately functioning microbial communities are also required to ensure that the produced sea-food is healthy and to prevent the proliferation of unwanted and pathogenic microorganisms. For example, about forty known

Figure 2.17. Recirculated fish tank with biofilter (drum on right), which uses beneficial microorganisms to remove chemical wastes from the water. (Courtesy Yonathan Zohar, COMB)

microalgal species produce potent toxins which, in the case of explosions of growth (HABs), may harm marine life and cause losses to aquaculture operations or a ban on sales of the products for human health considerations.

2.3.6 Human health

Several pathogenic organisms are ubiquitous to the marine environment. *Vibrio cholerae*, a brackish water microbe causes cholera, a disease that afflicts hundreds of thousands of people worldwide. About 200 serogroups of *V. cholerae* are recognized, but only serogroup O1 and the newly emerged O139 are associated with cholera pandemics. *V. cholerae* has been well-recognized for over 30 years to be endemic to aquatic environments, but the proliferation of virulent *V. cholerae* strains and the factors that affect disease dynamics in aquatic systems are still incompletely understood.

Places where cholera has been absent for decades have recently been plagued by cholera epidemics, particularly in West Africa. Riverine, estuarine and coastal waters are reservoirs for *V. cholerae*, harbouring virulence genes which may at any time revert to a readily transmissible infectious state under certain nutrient, pH, salinity and temperature conditions. The emergence of *V. cholerae* outbreaks in some cases has been correlated to increasing surface water temperature and global warming. Heating of oceanic surface waters, especially those near tropical or subtropical coasts, increases the abundance of copepods, a host for cholera bacteria. Also, in temperate estuaries, the abundance of *V. cholerae* in the environment is seasonal, and the association with zooplankton may provide some protection from the relatively harsh environmental conditions prevailing in aquatic habitats. In the temperate Baltic Sea, *Vibriovulnificus*, which can cause wound infections, gastroenteritis, or a syndrome known as primary septicaemia, was found to be responsible for some fatalities in immune-compromised persons.

2.3.7 Energy

One of the major societal challenges we face is the long-term supply of sustainable energy. Marine living sources can contribute importantly to secure the growing demand for energy in Europe and beyond. The two most promising approaches to harvest energy from marine living organisms are based on marine microbial life. The first approach, Microbial Enhanced Oil Recovery (MEOR), aims to utilize marine microbes to improve the recovery rates of classical fossil oil reserves. This is done by either decreasing the viscosity of oil and/or the permeability of the rock material in which the oil resides. The second one focuses on the production of biofuels from microalgae. Both of these applications require an in-depth knowledge of marine microbial diversity and functionality in order to select useful strains and/or to identify by-products that are of use in these production systems. At the time of writing, the application of MEOR and the production of biofuels from microalgae is either not yet applied, or not yet commercially viable and considerable research will be needed in the future to develop operational systems at a commercial scale which are controllable and environmentally sound. More details about the state-of-the-art and research challenges associated with these promising technologies can be found in the 2010 Marine Board-ESF Position Paper 15 on Marine Biotechnology.

2.4 Concluding remarks

Ninety per cent of the biomass in the ocean is microbial and represents an enormous living diversity. Relatively few types of microorganisms are abundant and control the major biogeochemical cycles. The vast majority are rare but represent an almost infinite genetic pool from which the microbial community can draw in order to be able to respond to environmental changes, thereby contributing to the stability of the ecosystem. The ocean also contains a huge number of viruses, exceeding that of *Bacteria* by an order of magnitude. Viruses may control microbial biomass and keep in check the successful microorganisms by 'killing the winner'. Viruses also generate and maintain biodiversity since they are the main mediators of genetic exchange between organisms. The balance of all these biogeochemical cycles and their compounds controls the dynamics of all ocean biomes. Thus, understanding the ecology of marine microbial

communities is essential for our understanding of the oceanic system, for the sustainable use of the ocean and to be able to anticipate changes in ecosystem functioning. A major problem that we face in studying changes in the ocean is that we are lacking a baseline of microbial ecosystems that we can refer to for assessing and quantifying the changes that we observe. In order to accomplish this it is not only necessary to make an inventory of which microorganisms currently live in the ocean, but also to reconstruct what has lived in the past ocean. Only then might we be able to predict what will live in the future ocean.

Fishing is a major human activity in the ocean as it supports a major part of the protein supply for humans. The production of fish depends mostly on primary production by phototrophic microorganisms, which, in turn, depends on the availability of nutrients. When nutrients are in excess it may lead to the development of HABs. Some of the species that produce HABs are toxic and may cause serious health problems. Even if these algae are not toxic they may reduce water quality and cause anoxia. Organic matter produced by phototrophic microorganisms and its transmission through the food web results in a huge amount of dissolved organic carbon (DOC) which is decomposed through the microbial food web, recycling nutrients which give rise to new biomass. The viral loop speeds up the recycling of nutrients and carbon. The microbial food web is responsible for the biological pump, namely the transport of CO_2 as organic carbon to the abyssal ocean and ocean floor. It is unclear how much CO_2 can be stored and what processes might return it to the atmosphere.

Nitrogen is a major driver of primary production but it is in short supply in vast areas of the ocean. Half of the world fixation of N_2 is anthropogenic (by the industrial Haber-Bosch process) and a large part is discharged in the marine environment through wet and dry deposition, river discharge, and run-off. Half of the natural global N_2 fixation takes place in the ocean representing an important driver of primary production and CO_2 fixation. N_2 fixation counteracts the losses of combined nitrogen through denitrification and anaerobic ammonium oxidation (anammox) although the contribution of the latter is uncertain.

Current knowledge of community processes at the genomic and biochemical level, taking into account the vast range of microbial interactions, is insufficient for the mechanistic and predictive understanding of the global CO_2, N_2O, DMS and CH_4 fluxes. These fluxes have an impact on the atmospheric composition and influence climate. Thus, while there is knowledge about the processes that mediate the emission of such climate change gases by microbes, it is unclear how these microbes influence climate. One of the reasons is the lack of understanding of the controls on the fluxes and concentrations of these gases in the surface ocean.

Marine microorganisms have an untold potential for providing services and products for human society which is not exploited to any significant extent. There is a world of microorganisms to discover, understand and put to good use.

Summary Box 2.1. Key research priorities and recommendations

1. Develop a census of marine microorganisms currently living in the ocean; develop a reconstruction of what has lived in the past ocean; and develop models which can predict what might live in the future ocean.
2. Produce a comprehensive inventory of all microbial metabolic pathways in the ocean including their genetic basis and analyze them in the context of marine microbial systems biology to better understand ecosystem processes and functioning.
3. Research is required to elucidate the environmental factors which control and regulate microbial metabolism and the ecological interactions (competition, cooperation, symbiosis) between different microorganisms, and how they influence ecosystem functioning.
4. Increase the efforts to understand the fluxes, sources and sinks of key components of the oceanic biogeochemical cycles through process-based measurements using innovative and novel technologies.
5. Increase the efforts to elucidate the CO_2 storage capacity of the ocean and initiate research to better understand the functioning of the biological pump and the processes that potentially return the stored CO_2 back to the atmosphere.
6. Increase the efforts to improve the understanding of the processes that control the efflux of the greenhouse gases CO_2, N_2O, DMS and CH_4 in the surface ocean.
7. Increase the efforts to explore and exploit marine microorganisms for biotechnological applications and products in line with the recommendations of Marine Board Position Paper 15 *Marine Biotechnology : A New Vision and Strategy for Europe**
8. Analyze oceanic metagenomic databases to identify putative new metabolic pathways and facilitate their discovery. In particular, target the understanding of signalling networks.
9. Drastically increase our knowledge of the biological, physical and chemical aspects of the deep (abyssal) ocean and ocean floor with relevance for the occurrence, diversity and functions of microorganisms.
10. Develop regulations and programmes for monitoring microbial changes as a result of human activities such as deep-sea mining.

* Marine Board Position Paper 15 *Marine Biotechnology: A New Vision and Strategy for Europe* is available for download from www.marineboard.eu/publications/

3. Technologies and research toolkits

3.1 Introduction

Microbial ecology is currently experiencing a renaissance spurred by the rapid development of molecular techniques and the so called omics technologies in particular. These tools have allowed researchers to produce a vast amount of information through *in situ* measurements and analysis of natural microbial communities. Such approaches are vital to unravel the interactions of microbes with their environment and with one another.

Today, genomics of DNA and RNA (transcriptomics), of proteins (proteomics) and of metabolites (metabolomics), used alone or in combination with each other and/or with more classical methods, are fields that are rapidly transforming many areas of biological research. In the future, however, it will be crucial to combine results obtained by application of these techniques together with information about the physical, chemical and other biological parameters in order to form a basis for addressing the societal questions raised in this position paper (see Chapter 1). Such multidisciplinary approaches are necessary to understand and potentially predict the complex role of marine microbial ecology in global change, sustainable supply of healthy food and in preserving human and environmental health. Likewise, information from these combined techniques integrated with the measurements of various abiotic and biotic factors, will give a better understanding of the complex role and interplay of marine microbial communities. Furthermore, such an integrative approach will almost certainly reveal a wealth of new metabolic processes and functions which could potentially support innovation in biotechnology and, therefore, for the bioeconomy as a whole.

Nevertheless, using cultivated microorganisms is still the only way to get detailed information about microbial characteristics and processes, thus highlighting the need to further focus on culturing microorganisms and developing better culturing techniques. This chapter provides an overview of the most important technologies and research toolkits available to marine microbiologists today, highlighting the critical barriers to progress.

3.2 Classical microbiology

3.2.1 Enrichment, isolation and cultivation

Traditionally, microbiology has relied mostly on the enrichment of microorganisms from their natural environment, followed by their isolation into pure culture allowing their growth and physiology to be studied. Enrichment media and conditions were selected based on the supposed physiology of the organism subject to study. But it was soon realised that this way of selection would probably enrich a microorganism with the desired basic physiological properties but not necessarily the organism that would prevail under natural conditions. For example, high starting concentrations of growth factors would select for an organism that grows at the highest maximum specific growth rate while in nature organisms often are competing for low substrate concentrations with low specific growth rates. In order to enrich microorganisms at low substrate concentrations and low specific growth rates, chemostats were applied using sterile growth media with a chosen limiting substrate.

To obtain pure (axenic) cultures, i.e. a clonal culture, an enriched microorganism is isolated and purified away from contaminating microorganisms. This is often done by plating the culture on a Petri dish in such a way that single cells are spatially separated, with the growth medium solidified with an agent, or in a tube when isolating anaerobic organisms. After incubation for a certain period under appropriate conditions, the single cells will have grown out to visible colonies (Figure 3.1).

Unfortunately, the approaches described above rarely result in the isolation of the dominant and presumably

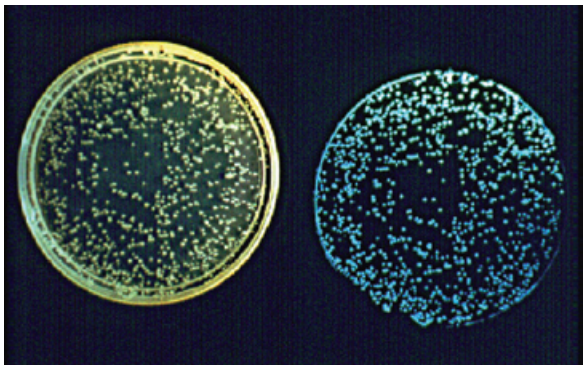


Figure 3.1. Colonies of the bioluminescent marine bacterium *Vibrio fischeri*. The photograph of colonies growing on agar (left) was taken with an artificial light source. The photograph of colonies on the right was taken using their own bioluminescence as a light source. (© J.W. Hastings, Harvard University, through E. G. Ruby, University of Hawaii; provided courtesy of the National Science Foundation)

ecologically most important microorganism from the environment. In addition, these approaches are time-consuming, in particular the chemostat approach, and other approaches have therefore been proposed. For instance, the direct plating of a diluted seawater sample on a Petri dish with a suitable growth medium would yield colonies of varying microorganisms that can be identified and further purified. However, this only yields those microorganisms that are capable of growing on the chosen medium and under the given incubation conditions. Another approach to just obtain dominant microorganisms is the dilution to extinction technique. A sample is diluted with sterile seawater from the same site as the sample from which the microorganisms are to be isolated until 1-10 cells are left.

Larger microorganisms (e.g. filamentous or large unicellular *Cyanobacteria*) can be directly isolated by micromanipulation. Using a finely drawn glass capillary and an appropriate microscope (e.g. a dissecting microscope), individual trichomes, cells or aggregates can be picked out from a community and transferred to a growth medium. A variety of other isolation techniques have been developed for specific microorganisms or samples. For example, the laser forceps technique in combination with microscopy in which a single cell is isolated from the rest. High throughput flow cytometric single-cell sorting (Figure 3.9) is another approach, applied to particles as small as viruses up to large phytoplankton cells such as *Cyanobacteria* or unicellular algae.

3.3 Novel culture-independent techniques

Despite having a large variety of different microbiological techniques at our disposal and the number of isolated and described microorganisms steadily increasing (31,000 *Bacteria* and *Archaea*), this number remains extremely low when compared to the estimated number of microorganism types, which stands at over 1 million. It is often stated, therefore, that we know less than 1% of the existing types of *Bacteria* and *Archaea*. However, that is not to say that more than 99% of the microorganisms which remain undescribed and uncultivated are by definition not culturable. In fact, it is more likely that we simply do not know how to isolate and grow these organisms for the following reasons. First of all, for many of the uncultivated microorganisms, there has been no serious attempt made to isolate them. Secondly, traditional techniques and procedures as described above are not enough for purifying and describing the remaining 99% in the coming decades, even if we knew how to isolate all types of microorganisms. This is because

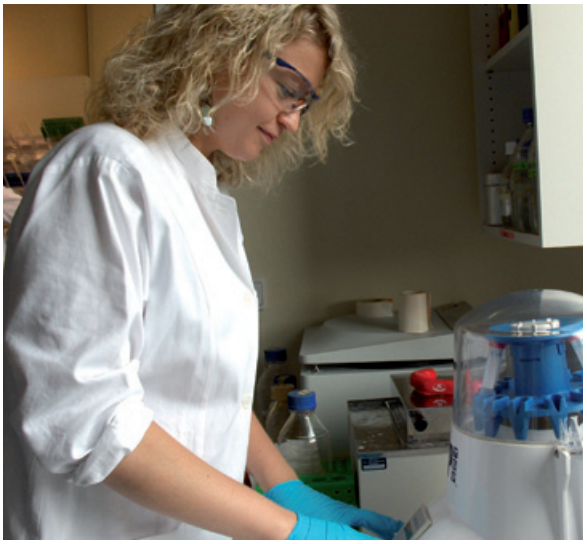


Figure 3.2. Marine microbiologist working in the laboratory (courtesy Frank Oliver Glöckner and Anna Klindworth)

they are too time-consuming and unsuitable for high throughput. Thirdly, it is possible that many, or even the majority, of the microorganisms are unable to grow as pure cultures in the laboratory. There are several examples of organisms that grow only in consortia with other microorganisms, even demonstrating typical multicellular behaviour and cell-to-cell communication.

Culturing new marine microorganisms will, however, always be of importance as it remains the only way to get detailed information about processes and functions from studies in controlled laboratory experiments. Culture-based techniques will also be the best approach for linking genes with function. However, cultivation of the vast majority of the uncultivated marine microorganisms introduces a challenge not only because most of the marine prokaryotes have not yet been cultivated but also because cultivation techniques often yield bacteria that are very rare in whole-community cloning studies. Culture-independent methods make it possible to study a whole microbial community and get a better understanding of the genetic diversity, population structure, and ecological roles of the majority of microorganisms in the marine environment. However, most of these culture–independent (molecular) studies tend to find primarily the most common organisms, like metagenomic sequencing that only (partly) enables reconstructions of the genomes of the most abundant members in the community. Thus, cultured bacterial and archaeal isolates that have yielded complete genome sequences will be important for evaluating metagenomic data. Recently, progress in the cultivation of microorganisms has accelerated, and developments in technology have

been a major factor in these advances (see above). A combination of methods now enables microbiologists to screen large numbers of cultures and to manipulate cells growing at low biomass densities that are characteristic of those found in the ocean.

3.3.1 Omics technologies

Marine microbial ecology is currently undergoing a paradigm shift, driven by the development and application of the so-called 'omics' technologies. Omics is a term used to denote approaches to analyze information embedded in available molecular subsets, hence, genomics, proteomics and metabolomics. Results from such studies have provided new and exciting insights in marine microbial ecology by adding new information about the complexity of the structure, interactions and functioning of marine microorganisms in the marine ecosystem. While genomics can provide information regarding the genetic potential of microbes, proteomics characterizes the primary end-stage product, proteins, thereby conveying functional information concerning microbial activity.

Metabolomics is the "systematic study of the unique chemical fingerprints that specific cellular processes leave behind". Specifically this means the study of their small-molecule metabolite profiles (Figure 3.3). The metabolome represents the collection of all metabolites in a biological organism or sample, which are the end products of cellular processes. Thus, while mRNA gene expression data and proteomic analyses do not tell the whole story of what might be happening in a cell, metabolic profiling can give an instantaneous snapshot of the physiology of a cell or given sample.

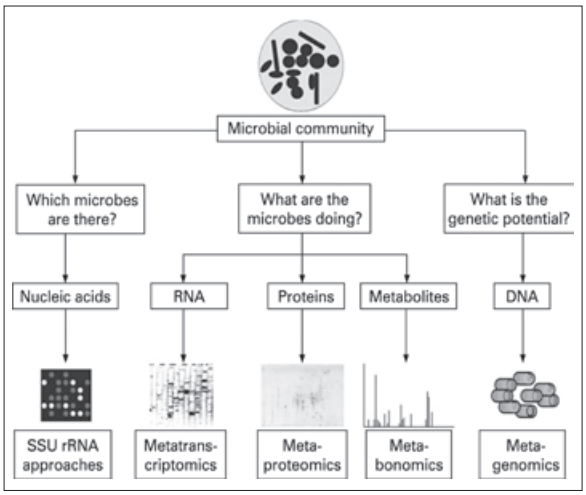


Figure 3.3. Schematic representation of community-based "omics" approaches. (from Zoetendal *et al*., 2008)

3.3.1.1 Genomics/metagenomics

Metagenomics is the culture-independent study of a community of microorganisms. Early environmental gene sequencing cloned specific genes (often the 16S or 18S rRNA gene) to produce a profile of diversity in a natural sample. Recent studies use "shotgun" Sanger sequencing or massively parallel pyrosequencing to get samples of all the genes from all members of the sampled communities. Advances in these technologies now permits two types of massive sequencing: (i) parallel pyrosequencing of amplicons from different samples, also called tag-sequencing, where the target gene often is part of the 16S rRNA gene or another core gene of functional or ecological importance; and (ii) high-throughput random sequencing of short gene products of total microbial community DNA or RNA.

The recent advances in metagenomics have revealed an enormous diversity of previously unknown, uncultured microorganisms that predominate in the ocean. However, these new techniques have also unravelled some new challenges that have to be focused on future research. One of the major challenges lies in the fact that most of the sequences so far identified show no similarity to previously described sequences, indicating an urgent need to develop novel research tools to characterise and describe the functions and ecology of the hereto unknown marine microorganisms.

3.3.1.2 Transcriptomics/metatranscriptomics

While the disciplines of genomics and metagenomics allow the study of the genomic potential of a particular organism or a microbial community, respectively, transcriptomics and metatranscriptomics deal with the subset of genes that are transcribed under certain environmental conditions. Accordingly, transcriptomics and meta- transcriptomics are powerful tools to capture snapshots of the genes essential for the survival of microorganisms under specific environmental conditions. Recent studies have shown the potential of such methods to provide detailed information on metabolic and biogeochemical responses of the microbial community to environmental change.

One of the unexpected results of the first metatranscriptomics studies was the detection of the high abundance and diversity of small RNAs (sRNA) in the marine microbial community. Microbial sRNAs are untranslated short transcripts that generally reside within intergenic regions on microbial genomes. Oceanic microbial sRNAs are thought to be involved in gene regulation and signalling processes related to environmental response. However, both the function and diversity of these genes are still unexplored.

Environmental transcriptomics protocols are technically difficult and a lot of effort has been given to develop

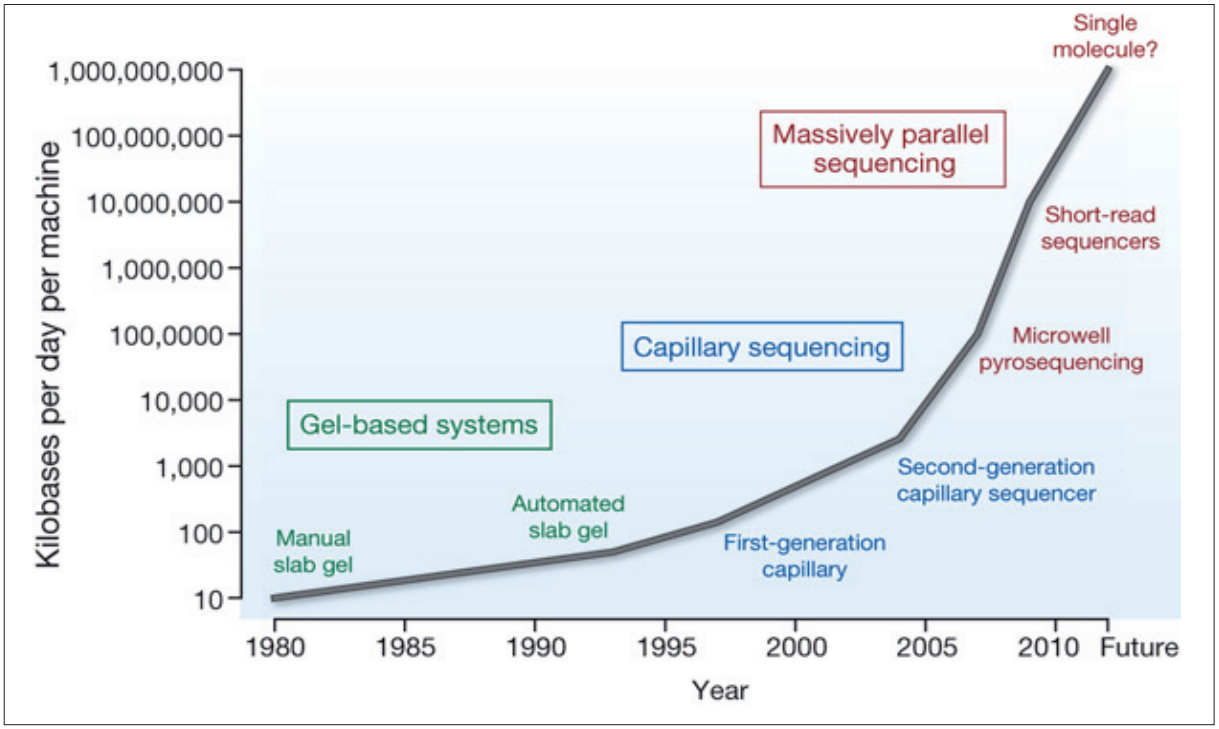


Figure 3.4. Improvements in the rate of DNA sequencing over the past 30 years and into the future.
(From MR Stratton *et al.*, 2009, Nature 458, 719-724)

new approaches that would produce reliable and unbiased data-sets. Further emphasis should be given to the development of new metatranscriptomic methods as this approach represents a powerful method for the study of metabolic processes and ecological activities within the microbial food-web. Such data, in conjunction with controlled field and laboratory experiments, will allow linkage of environmental variation with changes in the RNA pools and have the potential to provide new insights into environmental sensing and response (signalling) in natural microbial communities.

3.3.1.3 Proteomics/metaproteomics

Metaproteomics, in combination with metagenomics and bioinformatic binning methods, has the potential to provide a high-resolution representation of the active genotypic traits of distinct community members. Advances in mass spectrometry instrumentation and methodologies, along with bioinformatic approaches, has brought this analytic chemistry technique to relevance in the biological realm, thanks to its powerful applications in proteomics. Mass spectrometry-enabled proteomics, including “bottom-up” and “top-down” approaches, is capable of supplying a wealth of biologically-relevant information from simple protein cataloguing of the proteome of a microbial community to the identification of post-translational modifications of individual proteins. The techniques represent an emerging field of study which have been limited to a few studies so far, yet the potential of these techniques for the advancement of marine microbiology has already been clearly demonstrated.

All community proteomic methods face challenges due to the large complexity of protein species and the large dynamic range of protein levels. Current technologies can identify proteins of populations that comprise at least 1% of the community and for which closely related genomic sequences are available. Recent technological developments have improved the methods for extraction, separation, and identification of proteins from natural microbial communities, and have made this method both more sensitive and more accurate. However, these methods have so far been most suitable when applied to microbial communities with limited diversity and/or those dominated by particular organisms.

Metaproteomics has significant potential for solving one of the major challenges facing microbial ecologists in providing a high-resolution representation of community structure and function. The data generated will be important to provide crucial and fundamental information. Once integrated and analysed alongside metagenomic

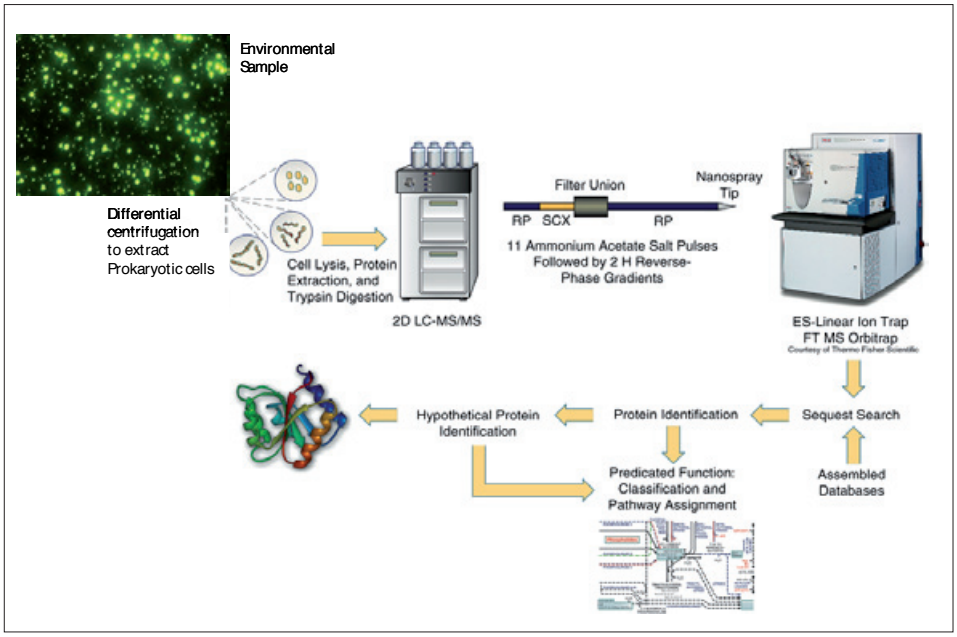


Figure 3.5. Shotgun metaproteomics approach used to identify microbial proteins environmental samples (Modified from Verberkmoes *et al.*, 2009)

data, metaproteomic datasets will provide information on the activity of genes and metabolic pathways that is currently not available.

3.3.1.4 Metabolomics

Metabolomics is an approach used to define the small-molecule diversity in the cell and to display differences in small molecule abundance. Metabolomics shows many advantages in terms of metabolic analyses because metabolites are the functional entities within the cells and their concentration levels vary as a consequence of genetic or physiological changes. The technique is one of the newest omics techniques and is, therefore, less developed and used in marine research. To date, most of the environmental metabolomics studies have been performed on higher eukaryotic organisms to measure biological, physical and chemical stressors and no studies have been reported on marine *Bacteria* or *Archaea*. In microbial ecology, the potential to employ the method will be limited by the difficulties in discriminating between metabolites of the three domains of life or even between the species within one of the domains. However, a promising approach is based on isotope-labelled intermediate metabolites combined with various types of dynamic metabolic flux modelling and may be a useful tool in investigations of complex and large-scale metabolic systems. Nevertheless, improvements are still needed in both instruments and data-analysing software to improve identification coverage and accuracy in order to make metabolomics compatible with other omics technologies.

3.3.2 Coupling identity with function

To associate the identity of a microbe with its function in the ecosystem there are several strategies. One of them is the so-called single-cell approach, by which we can use a series of probes to "interrogate" each cell at a time. One such probe is the "identity probe". Either before observation (by microscopy) or after some sort of physical separation (e.g. flow cytometry cell sorting), the probes are applied to the sample to collect information on its identity. Such probes include, for example, the FISH probes to the required phylogenetic depth level, but single-cell genomics screening with 16S rRNA PCR is also an option.

"Functional" probes can be of different types: (i) targeting the structure of the cell; and (ii) targeting the physiology of the cell (such as the redox probe CTC (5-cyano-2,3-ditolyl tetrazolium chloride) which indicates which cells are expressing an active electron chain, the probe ELF97 which indicates whether a cell is expressing phosphatase activity, or the probe PI (propidium iodide) which reports on the state of the membrane); or (iii) activity probes (such as radioactivity-labelled organic matter or nutrients that, combined to micro-autoradiography will indicate which cells are taking up the provided substrate (See also the section on Nano-SIMS below)).

In any case, the power of the analyses is made larger by combining different types of probes: identity plus functional probes. For example, techniques such as (i) stable-isotope probing, (ii) fluorescence *in situ* hybridi-

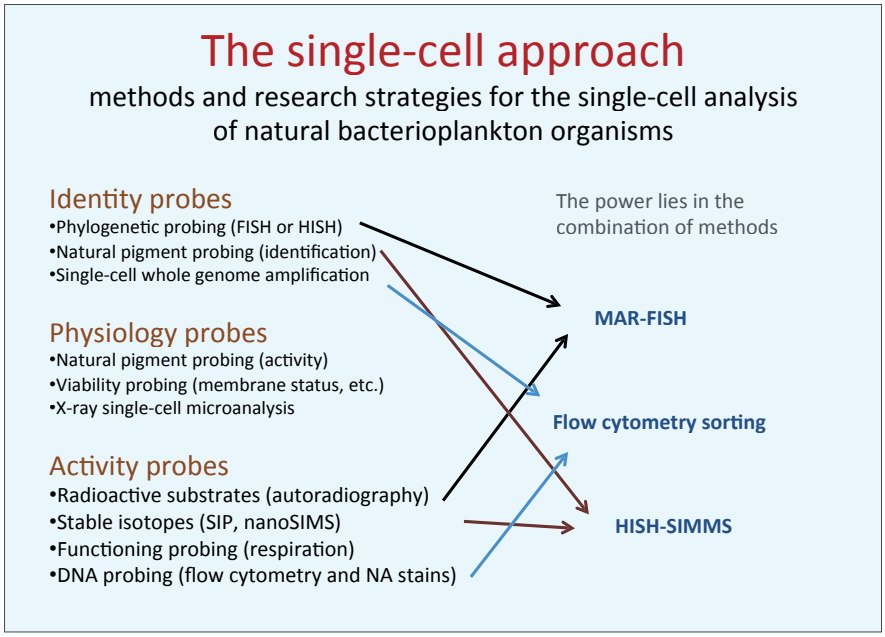


Figure 3.6. The single-cell approach

zation (FISH) micro-autoradiography (MAR-FISH), (iii) isotope array, or (iv) FISH-secondary-ion mass spectrometry (Nano-SIMS) will allow substrate incorporation and specific processes to be linked to microbial phylotypes. These single-cell techniques have been used to describe the environmental distribution and the functional role of microorganisms in soil, water, sediments and other ecosystems. The field has come a long way in the past decade to answer two basic ecological questions: "Who is there and how many? and What are they doing?"

3.3.2.1 Stable-isotope probing (SIP)

The goal of SIP lies in linking the identity of microbial communities with their function in the environment. The SIP technique involves introducing a stable isotope-labelled substrate into a microbial community and following the fate of the substrate by extracting molecular species such as fatty acids and nucleic acids from the community, determining which specific molecules have incorporated the isotope. Lipids and nucleic acids are used as the biomarker molecules. In SIP of lipids (PLFA-SIP) the approach is to expose a community of interest to an isotopically labelled substrate (usually ^{13}C or ^{15}N), then isolate lipids and trace the labels to specific molecules which can be characterized by gas chromatography. In nucleic acids-SIP (DNA- & RNA-SIP) the community is exposed to a labelled substrate (again usually ^{13}C). After labelling, either DNA or RNA is isolated and subjected to buoyant density gradient centrifugation. The nucleic acids of organisms that have assimilated the ^{13}C have a higher buoyant density than nucleic acids of organisms that have not. Following centrifugation, those nucleic acids that have taken up the substrate are amplified by PCR and its product may subsequently be identified by fingerprinting techniques or identified by sequencing. A major advantage of lipid-SIP is the high sensitivity, in contrast to nucleic acids-SIP where the synthesis of the labelled nucleic acid must be sufficient for the separation and detection of labelled molecules. A potential disadvantage of SIP studies is cross-feeding, requiring verification of the initial community members that consume a particular substrate and possible subsequent consumers of metabolic by-products, which has been seen as a potential problem for both DNA and RNA-SIP that require long incubation times.

All SIP applications are concerned with the detection of the assimilation of specific substrates. One example is to study the direct uptake of naturally occurring substrates by microorganisms or experiments to detect the flow of C into microorganisms. There are, of course, many possibilities for using this approach. Perhaps the most exciting studies are those which start with information from experiments of PLFA-SIP, where a particular group of microorganisms is involved in assimilation of a substrate. Following these studies, RNA-SIP from that group may be selected from the sample and subcloned, indicating which community members were assimilating the substrate. Future studies will likely include analysis with mRNA (mRNA-SIP) which would help in the study of the expression of genes from different environments.

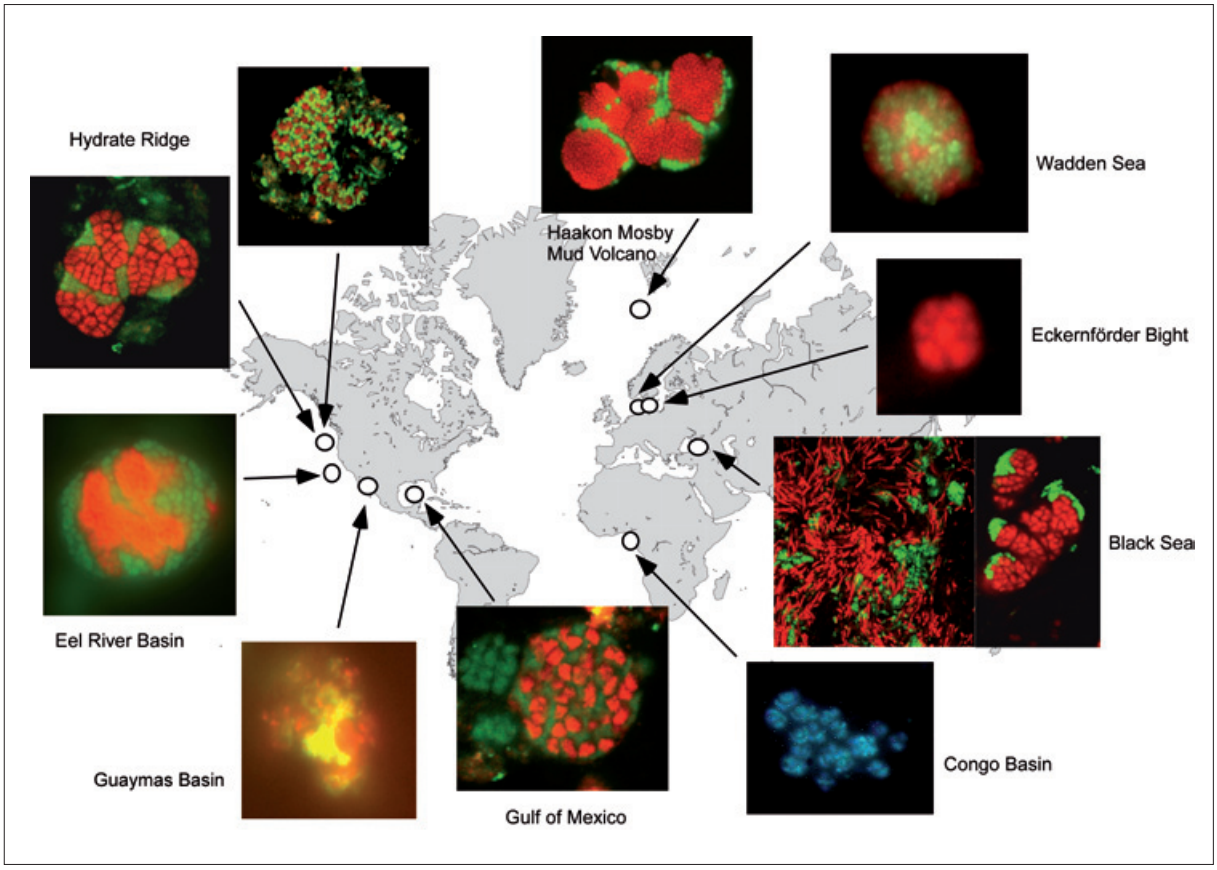


Figure 3.7. Different types of AOM (Anaerobic Oxidation of Methane) consortia visualized by fluorescence *in situ* hybridization (FISH). Red colours are ANME *Archaea* and green colours indicate sulfate-reducing bacteria. (Courtesy Dr. Katrin Knittel, MPI-Bremen)

3.3.2.2 Fluorescence in situ hybridization (FISH)-microautoradiography (MAR-FISH)

Microautoradiography is one of the earliest single-cell methods used in aquatic microbial ecology. In this technique, microbial assemblages are incubated with a radio-labelled substrate and cells are then placed in contact with an autoradiographic emulsion. Subsequent exposure of the emulsion to the radioactive emissions produces silver grain deposits around the cells that have incorporated the radioactive substrates. Acetate, glucose, amino acids were used in the initial studies, but recent studies have expanded to other substrates such as $^{14}CO_2$ or the carbon excreted by phytoplankton. It is possible to combine the microautoradiography technique with fluorescence *in situ* hybridization (MAR-FISH), providing information on the *in situ* single-cell activity of specific microbial groups. The application of MAR-FISH is limited to radio-isotopes with a suitable half-life and that can be obtained at high specific activities. In addition, the biologically relevant elements nitrogen and oxygen cannot be tracked by MAR-FISH, because of the lack of proper isotopes. However, MAR-FISH has the advantage that the active substrate uptake can be related to individual cells and that single-cell rates of substrate incorporation can be calculated. This is not the case with most methods that use stable isotope tracers.

3.3.2.3 Isotope array

Environmental arrays are gene fragments that are arrayed on microscopic grids which allow simultaneous monitoring of the diversity and substrate incorporation of complex microbial communities. DNA or RNA extracted from environments is hybridized with the gene array and the hybridization is measured as an indication of the presence of the active microorganism in the sample For instance, the diversity and bicarbonate incorporation of ammonia oxidizing bacteria (AOB) have been measured using a prototype DNA microarray for AOB detection in activated sludge samples. One advantage of microarrays compared to MAR-FISH is the ability to study the uptake of several substrates and apply many probes in

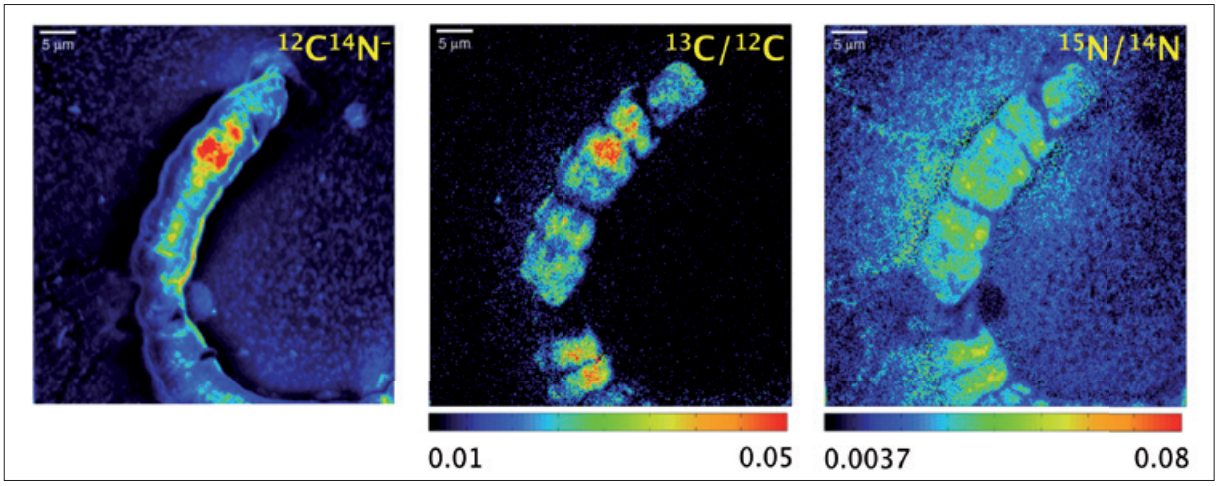


Figure 3.8. Enrichment of elements and their stable isotopes in a trichome of the cyanobacterium *Nodularia spumigena*. The left-side image shows the enrichment of carbon relative to nitrogen, indicating areas of active carbon fixation and storage. The middle image shows the accumulation of the stable isotope ^{13}C, relative to ^{12}C after feeding the sample with ^{13}C labelled bicarbonate, indicating the location of carbon fixation. The right-side image shows the same for nitrogen after supplying the organism with ^{15}N labelled N_2. (Picture courtesy of M. Kuypers)

parallel. Furthermore, the technique is neither radioactive nor toxic. However, the main limitation of the isotope array is its sensitivity and thus its application to samples collected from ecosystems with not very active or dense microbial communities, such as open ocean waters. In addition, substrate cross-feeding must be considered as it may cause problems in differentiating between primary substrate consumers and microorganisms which live on the lysis products of the primary consumers, particularly with prolonged substrate incubation time.

3.3.2.4 Secondary-ion mass spectrometry (Nano-SIMS)

Coupling the identity of microbes with their activity in the environment remains an important gap in our ability to explore microbial ecology. The development of techniques to quantify the metabolic activity of single microbial cells has been especially challenging, mostly due to their small size. A particularly promising technique is nanoscale Secondary Ion Mass Spectrometry (Nano-SIMS), a new high-resolution imaging method, which can help decipher what individual microbes are "doing" in the environment. High-resolution nanometer scale secondary-ion mass spectrometry (SIMS) has become established within environmental microbiology and this opens up new possibilities for the coupling of phylogenetic identity and metabolic function of single cells in studies of mixed microbial communities from the environment. Nano-SIMS is a surface analysis technique which provides information about the spatial distribution of any element and its isotopes as well as quantitative information about the isotopic composition of a sample. Nano-SIMS can determine the chemical, radioisotopic and stable-isotopic composition of biological material down to the submicrometer level. By exposing microbial communities to substrates labelled with isotopes, Nano-SIMS-based imaging allows visualization of metabolic activity in single cells. Moreover, substrate uptake rates and fluxes can be quantified.

The combination of Nano-SIMS with fluorescence *in situ* hybridization (FISH) finally provides the link between the identity of microbial cells and their activity. FISH uses fluorescent-labelled probes that are specific to the organism of interest and that bind to the intracellular 16S ribosomal RNA. The use of HRP-labelled oligonucleotide probes and catalyzed reported deposition of halogenated tyramides (HISH) makes the method very useful for the *in situ* analysis of a wide range of environmental samples. Replacing fluorescent probes with halogenated probes allows individual cells to be directly identified (by probe hybridization to targets) by Nano-SIMS.

However, one of the disadvantages of Nano-SIMS technology which limits its widespread use is the expense of the equipment. Alternatives such as RAMAN spectroscopy which is much cheaper and still allowing functional information of single cells are, therefore, noteworthy. In RAMAN spectroscopy, a spectroscopic technique is used to study vibrational, rotational, and other low-frequency modes in a system. It relies on inelastic scattering, or Raman scattering, of monochromatic light, usually from a laser in the visible, near infrared, or near ultraviolet range.

3.4 Why models?

The obvious challenge to the scientific community of marine microbial ecologists is to transform the avalanche of omics data at the molecular scale into "understanding" at the time and space scales of marine ecosystems. In a traditional hypothetical-deductive perspective, science is seen as a spiralling process where observations lead to models (insights, hypotheses) that can be used to construct new experiments that generate new data, and where each turn brings us a step up along the spiral towards a better "understanding". In this generalized context where "models" are the conceptual frameworks we construct to relate observations and make predictions, models are an integral and obligatory part of the research process and data collection without models remains just data collection, just as models without any root in data remain just models. Although the possibility for automating the whole process, hypothesis generation included, has been discussed (King *et al.* 2004) the need to extract meaningful "understanding" from huge sets of data is likely to increase the need for models.

Looking back at the last decades of marine microbial ecology, an important refinement to this picture becomes obvious: Journeys into new and unknown territories requires open-minded and model-independent observations to allow an unbiased observer to be able to see "the unexpected" and thus create a new entrance into the hypothetical-deductive spiral. You need to know that there are ~10^9 bacteria and ~10^{10} virus L^{-1} seawater before you can start looking for mechanisms behind and relationships between such data.

A characteristic trait of a mature field of science is a common theoretical framework within which observations can be organized and related, new observations predicted, and therefore experiments planned to challenge the theory. Despite its progress, one can ask whether marine microbiology scores particularly well on such a criterion. Perhaps is it still to a large extent in the initial descriptive phase where it gathers increasingly detailed descriptions of the unknown? There is a huge effort to measure microbial diversity, but to what extent is this still the initial hunt for a correct number, more than a hypothesis-driven search for the underlying mechanisms and relationships producing a particular diversity? The abundance of viruses is well known, but there is no satisfactory theory explaining the order of magnitude of this number. Meta-analyses can generate correlation plots between bacterial production and primary production, but to what extent can these correlations be matched with mechanistic explanations for how the system works? Is the correct description of the 'state-of-the-art' that we have a lot of observations and data on what the marine microbial ecosystem looks like in terms of numbers, fluxes and even species, strains and genes, but still a relatively vague idea of the mechanisms relating these numbers? From this perspective, the new wave of molecular data is flooding into a field that is immature in the sense that it does not yet have a mature theory within which these datasets can be arranged, analysed and interpreted.

One optimistic argument is that this is exactly what is needed: with all this new data, there will finally be "all" the pieces of the puzzle, and the "modeling" part of the job, putting the pieces together, will then be trivial (or at least easier). Such optimism is not necessarily supported by the history of science. Compared to explaining microbial ecology from molecular data, the task of deriving heliocentric models for planetary motion from observing the night sky may (in hindsight) seem simple, yet required an observational period lasting through the history of mankind up to Copernicus and Kepler. Relevant processes in microbial ecology cover a continuous spectrum from molecular reactions via cells and food webs to global element cycles. For many of us, the insight that climate is related to nanoscale microbial interactions is part of our fascination for this field. However, detailed low-level descriptions are required to understand the systems' behaviour at higher levels. The traditional example is classical and statistical thermodynamics where factors such as temperature, pressure, volume and entropy could be related to each other without the later insight that they represent average properties based on the behaviour of a vast number of colliding gas molecules. Since the essence of conceptual modelling is to "explain as much as possible with as little as possible", an overwhelming amount of detailed molecular data may even create situations where "one cannot see the forest because of all the beautiful trees."

In conclusion, a maturation of the field of marine microbial ecology is being hampered by a lack of "models"; both those relating entities at the community level (phytoplankton, protists, prokaryotes, viruses), at the cellular and molecular levels and, in particular, "models" connecting such levels.

3.5 Expected future developments in marine microbial cultivation and -omics studies

Understanding the complexity of the structure, interaction and functioning of microorganisms in the marine ecosystem, is the key to understand and thus to protect and sustain marine ecosystems. This understanding is important for addressing the societal and scientific questions pointed out in this position paper. In order to

achieve progress in the quest for a better scientific basis for action, new developments in both classical and novel techniques are essential. No single omics analysis will fully unravel the complexity of fundamental microbial biology. Therefore, integration of multiple layers of information through multi-omics approaches in combination with environmental data layers (e.g. environmental data streams stemming from sampling, ocean observatories and remote sensing) and classical techniques is required. By applying such approaches one will eventually gain a more precise picture of the function, interaction and diversity within the microbial food-web.

As a consequence of the rapid rate of discovery in the omics era, there will probably be a resurgence in traditional ecological and functional approaches that will be essential in order to place our new discoveries into their proper environmental context. One example is the development of genomics of uncultured single cells by combining flow cytometry and sequencing (single-cell genomics; see below). Another example of mutually beneficial combinations of classical and novel technology is the combination of metagenomic sequencing with stable isotopic analysis of geochemical signatures. The latter method provides information both about the identity and interplay between microorganisms and chemical processes within microbial communities.

The single cell genomics method is based on the combination of cell sorting by flow cytometry and whole genome amplification (WGA), followed by sequencing (Figure 3.9). By using the DNA from single cells it is thus possible to generate reference genomes of uncultured taxa from a complex microbial community of marine microorganisms. Thus, the method allows for the description of the functional diversity within the numerically significant uncultured fraction of the marine microbial community, avoiding the difficulties associated with cultivation. Additionally, the information can also be beneficial to developing and improving culture techniques of the so far uncultured fraction of the microbial community.

One limitation which is intrinsic to metagenome datasets is the difficulty with linking genes to the organisms carrying the genes. Single cell techniques might be useful in

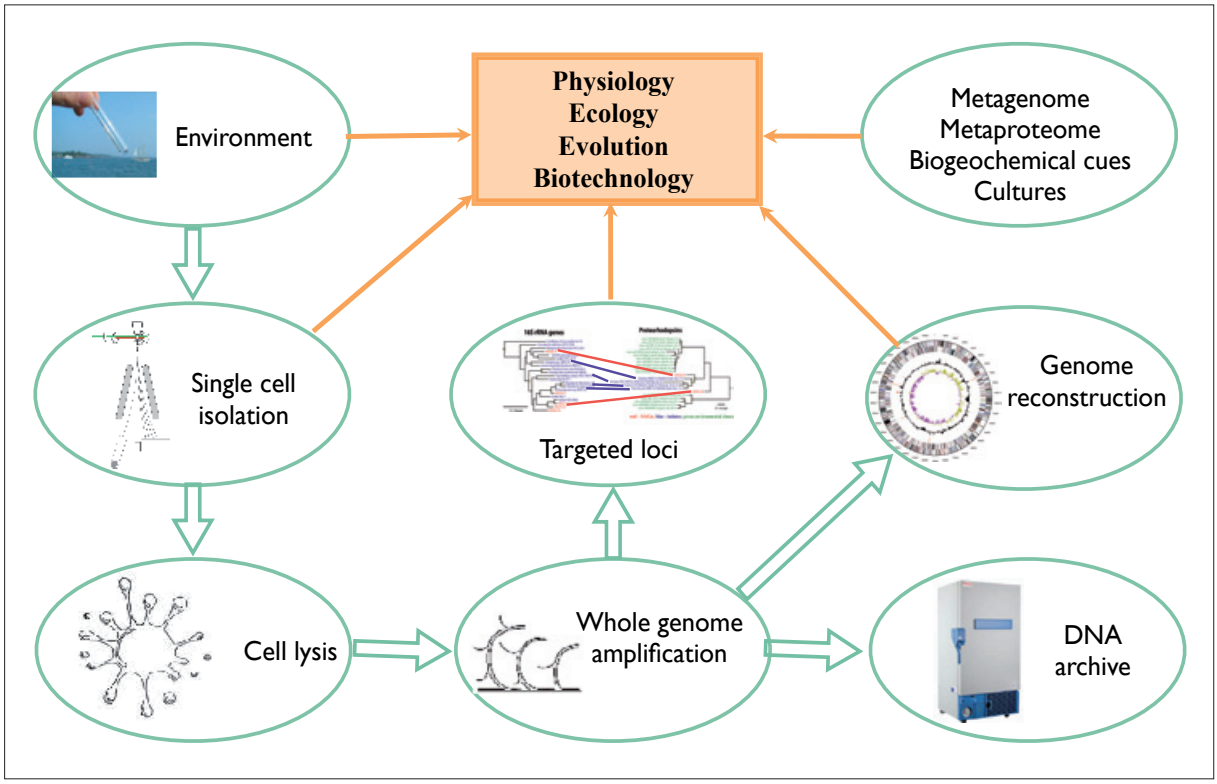


Figure 3.9. The single cell genomics pipeline: flow cytometry and sorting combined with single cell genomics. The WGA technique is a non-PCR based DNA amplification technique that has proven efficient in the amplification of small amounts of DNA, including DNA from single cells, resulting in a reasonable quantity for genomic analysis. Another alternative also includes flow cytometry cell sorting, but based on the natural fluorescence of some microorganism (Figure by Ramunas Stepanauskas)

overcoming this limitation. One example is data from the GOS expedition showing a predominance of photorhodopsin genes in marine metagenome libraries. The gene encodes a function of great importance to the marine food web, as the gene is potentially involved in the light-driven energy flux in ocean ecosystems. The dataset, however, cannot be used to assign the photorhodopsin gene to any specific organism. By combination of single cell genomics and metagenomics it will be possible in future research to analyze the genome content, metabolic adaptation and biogeography of specific groups of uncultured microorganisms and connect this information to metabolic pathways, ecological niches and evolutionary histories of marine microorganisms that are important players in the marine microbial food-web.

With the omic technologies, it has proved difficult to derive useful information about functions as so much of the sequence information has no database matches. In the future, unravelling the functions of all these sequences will be of great importance for expanding the knowledge about microbial processes and function in the marine microbial food-web. Future advances involving a combination of tools, both classical and novel, might be useful for uncovering the functional role of these unknown genes. This might include a combination of methods such as microarrays, metatranscriptomic, and proteomics together with metagenomics. Ultimately though, characterization of unknown genes requires the integration of classical methods, e.g. biochemical assay of enzyme activity or characterization of gene knockouts.

To date, all omics studies at the metascale have been carried out on bulk samples. However, microbial communities that exhibit distinct organismal and functional organization and particular enzyme variants may be localized within distinct microniches. Future research will need to resolve the functional significance of gene and protein localization within microbial communities. Hence, more advancement in fine-scale measurements needs to be developed. Additionally, more knowledge of low abundance community members is required, as these organisms have been overlooked by present technologies. New low-cost, high-throughput sequencing techniques and the technical developments in proteomics will increase the range of detectability both at the protein and nucleic acid level. These improvements will, in turn, increase the possibility of discovering the distributions and functions of rare organisms.

Advancements in marine microbiology have been traditionally driven by progress in other disciplines such as medicine and biotechnology. Marine microbiologists have not traditionally been trained to think about the development of technological solutions for specific problems. This is an issue that needs to be addressed for future improvements within the field of marine microbial ecology. One example exists, however, where medicine has benefited from both methodological development and knowledge present in marine microbial ecology. Community-wide sequencing (metagenomics) of viruses and prokaryotes in the marine environment has provided tools that have recently been applied to human health. Recent studies have investigated the viral community in human respiratory tract and faeces and have provided important early insight into these systems. The viral communities found in hosts with cystic fibrosis have been shown to differ greatly from communities in a healthy host, suggesting that viruses may play a major role in bacterial population control in the human oral cavity and respiratory system. Thus, not only the methodology but also the present microbial ecology knowledge may be important for understanding the processes and interactions between microorganisms in other disciplines.

3.6 Concluding remarks

In order to achieve progress in the quest for a better scientific basis to take measures and address the societal and scientific questions pointed out in this position paper, there is a need for development in both classical and novel techniques. For the future, it will be crucial to combine results obtained by the application of different techniques together with information combined with abiotic and biotic information. Integration of multiple layers of information through the multi-omics approaches, with classical methods, novel microscopy techniques and environmental data is required. By applying such an approach it will eventually be possible to get a more precise picture of the function, interaction and diversity within microbial foodwebs. There is a fundamental need to focus on developing technology that is specifically designed for studying marine microbial ecology.

Summary Box 3.1. Key research priorities and recommendations

1. Increase efforts to develop technological solutions for specific marine microbial ecology problems. Traditionally, advancements in marine microbiology have been mostly driven by progress in other disciplines e.g. medicine and biotechnology.
2. Strengthen the focus on culturing microorganisms and developing better cultivation techniques as information from microorganisms in cultures is still the best way to get detailed information about processes and functions.
3. Advance single cell methods enabling the coupling of microbe identity with activity.
4. Develop more low-cost, high-throughput sequencing techniques and technical developments in proteomics which would increase the range of detectability both at the protein and nucleic acid level. More knowledge about low abundance community members needs to be collected, as these organisms have been overlooked by current technologies.
5. Develop faster methods to unravel the functions of unknown genes. With the omics technologies it has been difficult to derive useful information about functions since much of the sequence information has no database matches. In the future, this will be of great importance for expanding the knowledge about microbial processes and functions in the marine microbial food web.
6. Improve bioinformatics tools to analyse genetic sequences.
7. Stimulate advancement in fine-scale measurements and process-based measurements. Microbial communities exhibit distinct organism and functional organization, and particular enzyme variants may be localized within distinct microniches. Future research has to resolve the functional significance of gene and protein localization within microbial communities.
8. Ensure a better integration of genomics, proteomic, transcriptomic, and metabolomic information together with physical, chemical and biological parameters to address fundamental smaller-scale scientific questions in areas such as microbial ecology, biodiversity and evolution. To further overcome current methodological limitations of the separate omics technologies, attention should also be given to combining these technologies both with each other and with traditional classical and novel techniques.

4. Marine microbiology observation and data infrastructures

Marine microbial researchers are mostly interested in obtaining the following two main types of data: (i) data about microbe diversity; and (ii) data about ecosystem functions driven by microbes. The tools and techniques used by marine microbiologists for sampling, data collection, management and analysis, are similar to those used by other marine disciplines. However, the study of microorganisms poses specific problems, which require additional, specialized procedures, techniques and tools.

This chapter provides an overview of the current status and practices in sampling and data collection processes within the marine microbial research community. It describes the process of managing the results of marine microbial research (data, information products, publications), and, in particular, marine microbial research products (samples, cultures and sequences) which we will collectively call Microbial Deliverables (MDs) for the purpose of this position paper. These MDs can be considered the raw data products generated by any research program, but they are not necessarily the *final* goal of research, which might be to understand the flow of carbon or nutrients, the production of an antibiotic product, the biodiversity patterns in a series of environments, or to isolate a not-yet cultivated organism etc. MDs can be considered the raw data products generated by any research program. Finally, this chapter also identifies key technological and organizational gaps, and formulates possible solutions and recommendations to improve future research.

4.1 Types of microbial deliverables

Marine microbiology research generates products which can be assigned to one of the following types (see also Figure 4.1):

(i) Geographic and oceanographic data
Environmental descriptors are often referred to as "ancillary data" in studies of marine microbial diversity. This includes purely geographic data (position, depth of sample origin), oceanographic (temperature, salinity, density, etc.), chemical (oxygen, inorganic and/or organic nutrient concentrations) or biological (chlorophyll, prokaryote abundance) data, etc. Data can also be split into state variables (concentrations, abundances) and rates (leucine incorporation, nitrogen fixation, etc.). Some of these data are obtained while sampling via the ship GPS positioning systems, or with a CTD (conductivity, temperature and depth) probe or a cluster of sensors and Niskin bottle samplers attached to a rosette. Others require further laboratory work to obtain the precise measurement of e.g. chlorophyll a, prokaryote abundance or bacterial production.

(ii) Sorted and preserved isolates, cultures of isolates (in liquid or solid media), or single-cells.
In general, long-term maintenance can only be assured by sample freezing.

(iii) Concentrates of microbial biomass
Microbial biomass concentrates can be obtained with a specific prefiltration step and a specific filter type and pore size. The concentrates include purified DNA, RNA, proteins and the rest of the metabolome along with cell walls and other cellular fragments. The concentrates might have the DNA or RNA extracted (DNA or RNA extracts), and the RNA might be converted to cDNA as this is more stable and the subject of most subsequent molecular analyses. It is often the case that laboratories keep a stock of the untreated original sample concentrate, another one with the nucleic acids extracted, and a third one with the cDNA products. All of these are commonly kept frozen. When cloning is used, the *E. coli* clones themselves are commonly stored for a given amount of time.

(iv) Genetic sequences
Genetic sequences may result from (a) metagenomic studies in which no PCR amplification has been used; (b) PCR amplification of a single gene (e.g. 16S rRNA or functional genes such as *pufM*). Data can originate from either Sanger (typically sequences of >500 bp) or short-read (60-400 bp) sequences from 454 pyrosequencing, Illumina, Solexa, SOLiD, etc. Additionally, proteomic studies produce protein

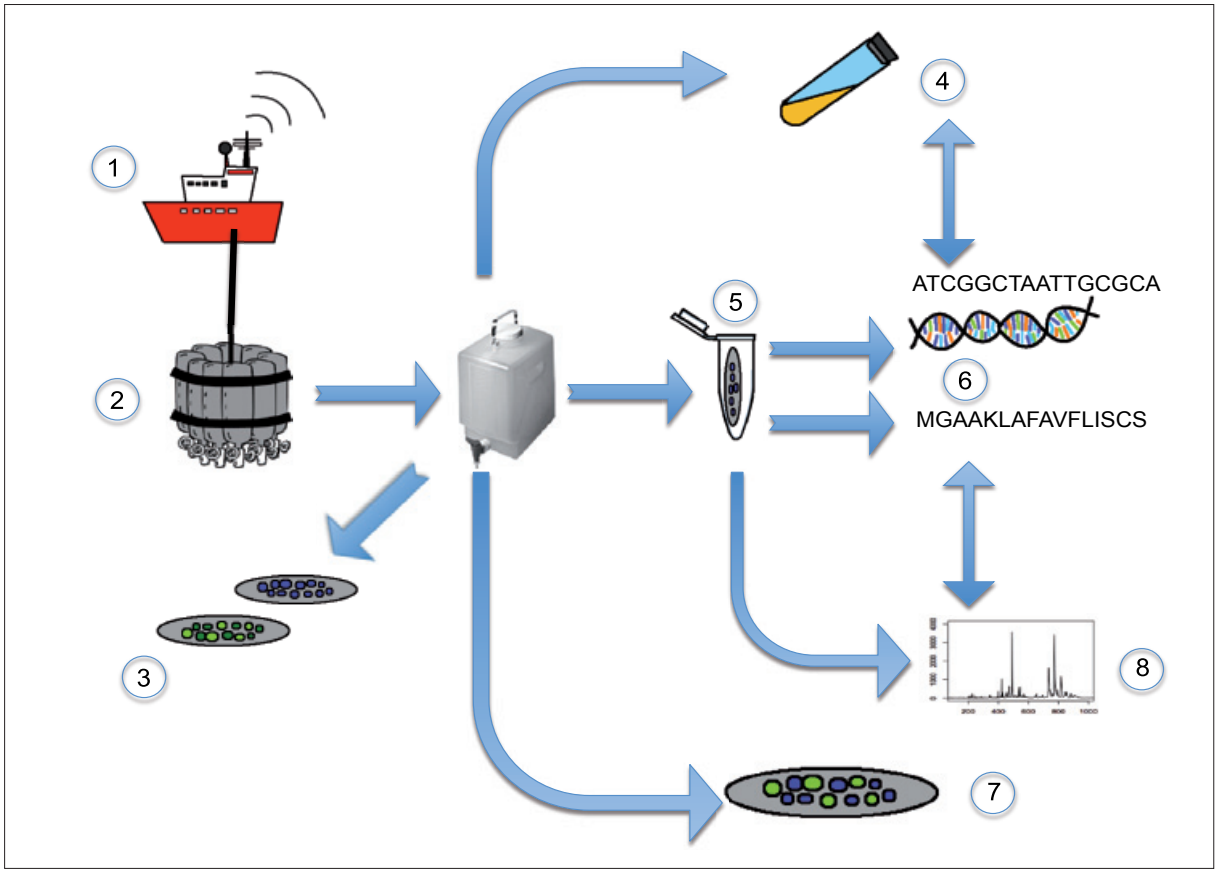


Figure 4.1. Types of data in studies of marine microbial diversity and function: (1) Geographic data; (2) Oceanographic data collected by the CTD sensors; (3) Chemical and Biological ancillary data; (4) Isolates from a given sample; (5) Microbial biomass or DNA stocks; (6) DNA, RNA and/or protein sequences; (7) Fluorescence In Situ Hybridization (FISH) samples; (8) fingerprints of the distribution of different sequences of a specific gene. (Image by C. Ruiz-González and J.M. Gasol)

sequences either directly (e.g. ESI-MS: electron-spray ionization mass spectrometry) or after a gel electrophoresis step. When using PCR amplification, the specific conditions used (primers, chemistry, machine type, conditions) should somewhere be explicitly stated. Also, it should be clear whether the data have been cleaned or treated in any way, or whether they are raw untreated data.

(v) Fingerprinting analyses
Fingerprinting analyses (e.g. ARISA, T-RFLP, DGGE) are often performed with the amplified DNA of marker genes like the ribosomal RNA. However, a lack of standardization in specific analysis methods between laboratories means that obtained gels and datasets are not always documented or linked to the rest of the data.

(vi) Original samples
Water samples can be filtered and stored to perform subsequent quantitative single cell counting e.g. by Fluorescence In Situ Hybridization (FISH). In these cases the samples might be stored before or after hybridization and/or after the results are recorded.

Ideally, data from the different types listed above should be integrated so that, for example, an isolate can be linked to a specific sampling position and corresponding oceanographic and biological conditions and to the isolate's rRNA and other (meta)genomic data. Figure 4.1 illustrates these different types of data and the flow between them.

The abovementioned products are obtained from three basic types of studies which are further discussed in section 4.2. These are
(i) Samples of opportunity;
(ii) Planned oceanographic cruises; and
(iii) Fixed time-series studies.

4.2 Origin of microbiological data

4.2.1 Samples of opportunity

Samples of opportunity are those samples taken during a specific cruise or trip organized for a purpose other than research. Sometimes such samples are collected for microbial analysis.

Samples of opportunity have been very important in the past. They have often generated the most relevant changes of paradigms because they tend to be used to test the possibilities of new technologies and methodologies. In general, new equipment and techniques are applied only to a limited number of samples because they are often expensive and time-consuming. As an example, the first metagenomics studies were done with samples collected in an isolated way, without any study of the surrounding physical, biological or ecological environment. Samples of opportunity are sometimes collected from a sailboat with a bucket with no measurements of chlorophyll or temperature to accompany the data. This lack of concurrent determination of environmental parameters certainly limits the possibilities of posterior interpretation of the data. It is therefore recommended that after the initial "test-of-concept" of the methodologies, this type of sampling should be avoided.

4.2.2 Oceanographic cruises

Oceanographic cruises are funded and undertaken specifically for scientific research. Unfortunately, few oceanographic cruises are purely dedicated to microbial research. Shiptime is expensive and often has to be shared with several other research groups. As a result, most research goals have to be agreed upon by the participant crew and samples are not always taken at the times and sites preferred by microbial researchers. Furthermore, cruises tend to be of "one occasion only" (e.g. Mediterranean cruise BOUM[1]). Specific exceptions to this trend are cruises such as the Atlantic Meridional Transect (AMT) which runs through the central Atlantic Ocean annually, or the series of cruises that have explored with similar methodologies the deep waters of the central Atlantic (cruises TRANSAT, ARCHIMEDES, etc.[2]). In all these cases, however, the spatial component of microbial diversity tends to be studied in more detail than the temporal component.

While there are documents that explicitly describe the consensus of the oceanographic community in terms of unit usage in chemical oceanography and in biogeochemical research[3], there is currently no agreement

Figure 4.2. Examples of oceanographic ships used in scientific cruises: (a) Hespérides (Spain); (b) Polarstern (Germany); (c) Oceania (Poland); (d) Sorcerer at Helgoland Roads (Germany); (e) Pourquoi pas? (France); (f) Pelagia (The Netherlands). (From upper left to down right: courtesy J.M. Gasol, AWI, Euroceans, Frank Oliver Glöckner, Ifremer, NIOZ)

on how samples should be collected and treated, and which ancillary information should be collected alongside the samples that will undergo molecular analyses for diversity studies. It is suggested that recommendations should be set for the collection and storage of marine microbial samples, taking into account the existing recommendations in place for genome analysis and characteristic determination. Similarly, homogeneity in how ancillary data should be collected, processed and stored would be desirable. In a recent effort the Genomic Standards Consortium (www.gensc.org) defined metadata checklists for genomic, metagenomic and marker genes sequences for a set of environments which can guide the collection and storage of ancillary data.

1. http://www.com.univ-mrs.fr/BOUM/
2. See also http://www.microbial-oceanography.eu/expeditions.html
3. UNESCO (1985) The international system of units (SI) in oceanography, UNESCO Technical Papers No. 45, IAPSO Pub. Sci. No. 32, Paris, France, and UNESCO (1994) Protocols for the Joint Global Ocean Flux Study (JGOFS) core measurements. IOC Manuals and Guides, n. 29. 170 pp.

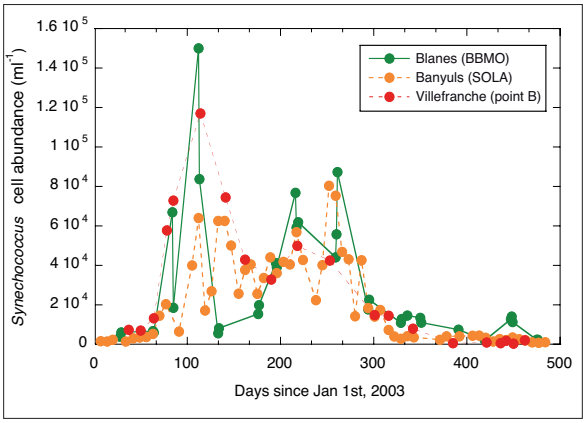


Figure 4.3. Synchronicity in the development of populations of a group of microbes (*Cyanobacteria* of the genus *Synechococcus*) during 2003-2004 at three microbial observatories in the NW Mediterranean separated by a total of ca 400 km: Blanes Bay, station MOLA at Banyuls-sur-mer, and Point B at Villefranche-sur-mer. Note the strong differences in spring and the similarities in summer and autumn (unpublished data of P. Lebaron, M. Weinbauer and J.M. Gasol).

4.2.3 Time-series and microbial observatories

Fixed time-series stations sample the marine environment regularly in time. Fixed time-series observatories for monitoring microbial diversity are predominantly coastal, although some important open ocean stations exist. For instance, two time-series stations initially focused on marine biogeochemistry through the international JGOFS[4] project, the HOT and BATS stations[5] in the North Pacific and North Atlantic respectively, have been fundamental in our furthering knowledge of the magnitude and dynamics of microbes and microbial processes in the ocean. However, other biogeochemically-focused open-ocean time-series stations do not have a strong biological microbial diversity component (e.g. station ESTOC[6] near the Canary Islands or VACLAN[7] in the deep North Atlantic waters). Except for a Mediterranean station (DYFAMED[8]), the rest of European stations with a component of microbial diversity are coastal, such as stations Blanes Bay, MOLA and SOLA in the NW Mediterranean region, Gulf of Naples, Helgoland roads and L4 in the North Sea and English Channel, and the RADIALES[9] in NE Iberian peninsula, etc. These stations typically run their programs in a largely uncoordinated way with the exception of EU-funded projects (e.g. project BASICS and the Blanes Bay, MOLA and Villefranche-point B stations, see Figure 4.3). This is a serious limitation to their utility. If all the stations were to use the same methodologies, analyze the same ancillary variables, and coordinate their efforts, we would gain a much better understanding of the seasonal variability of microbial diversity and function in the oceans surrounding Europe. Such integration activities have been funded for some key open ocean observatories across Europe (e.g. EuroSITES, ESONET) but these do not yet have strong microbial diversity components.

4.2.4 Automatic data acquisition technologies

The wide-spread use of automatic sampling for microbial diversity would certainly be desirable, but we are still far from having at hand the technology needed for such an endeavour. While sampling is relatively time-consuming when large amounts of water need to be processed e.g. in serial filtrations to account for organisms of different size classes, processing of the samples is particularly time-consuming. Firstly, processing requires several series of filtration, DNA/RNA extraction, PCR amplification, etc., and secondly, it is difficult to technically automate the protocol. Two possibilities that are under development and worth mentioning are the *in situ* flow cytometers and the Environmental Sampling Processor (ESP) (see also Chapter 3).

In situ flow cytometers take samples at regular intervals, sometimes staining or fixing the samples, and run them through one or several UV, blue or red lasers to identify microbes either by their size, their morphology or their DNA content. Some machines can be submerged and operate on their own for months, reporting back to the main lab by radio or other type of signalling. They are already in operation in various places in the US, and a company in Europe is producing several types of such machines[10]. Some instruments (e.g. FlowCytobot in Gulf of Mexico waters), have been able to produce data that has allowed early warning of the appearance of toxic dinoflagellates. They can still be considered, however, in the development phase of most of the technologies.

The ESP goes a step further. It is currently a prototype created by the MBARI (Monterey Bay Aquarium Research Institute) that provides on-site (*in situ*) collection and analysis of water samples from the subsurface ocean. The instrument is an electromechanical/fluidic system designed to collect discrete water samples, concentrate microorganisms or particles, and automate application of molecular probes which identify microorganisms and their gene products. The ESP also archives

4. http://www1.whoi.edu/
5. http://hahana.soest.hawaii.edu/hot/
6. http://www.estoc.es/en/
7. http://www.vaclan-ieo.es/index_en.html.
8. http://outreach.eurosites.info/outreach/DeepOceans/station.php?id=2
9. http://www.seriestemporales-ieo.com/en/index.htm
10. www.cytobuoy.com

samples so that further analyses may be done after the instrument is recovered. The machine has found success in the detection of dinoflagellate harmful algal bloom species, and in the detection of some specific bacteria and other organisms by means of probe microarrays. Microfluidics systems that would allow a PCR reaction followed by capillary electrophoresis are being designed. No such technology is being produced or assayed in Europe for marine research.

Relevant to instruments such as the one above, the term "ecogenomic sensing" has been coined with the idea of detecting molecular markers indicative of specific organisms, genes or other biomarkers within an environmental context. A field-portable device applied towards oceanographic research or water quality monitoring using wet chemistry molecular analytical techniques to assess the presence and abundance of a specified set of organisms, their genes and/or metabolites in near real-time is the objective behind the ESP (Figure 4.4).

4.2.5 Experimental manipulations

Data often originate from controlled manipulation experiments investigating marine microbial functioning. Examples include mesoscosm experiments that have tested the effect of nutrients on microbial dynamics, or those that test for the effects of global change (temperature increases, acidification), or the addition of pollutants (e.g. black carbon and oil, etc.) or natural products (aerosols, Sahara dust, etc.). Some of these experiments have been performed in large-scale tanks in the open ocean, extreme cases being the iron addition experiments (e.g. SEEDS, EiFEX, SOIREE, SOFEX, SAGE…) and the phosphorus *in situ* addition experiment (Eastern Mediterranean Sea, CYCLOPS project). Manipulations might also include the removal of predators (cascading experiments) or many other biological forcings. In these cases it is likely that the sample origin is the same for all treatments, but that the main factor differing between samples is the treatment. It is extremely important that the data that come from these types of experiments are correctly labelled as coming from an experiment, and with labels that indicate the treatment. In that way, the development of e.g. *Alcanivorax* and *Cycloclasticus* can be associated to the presence of added hydrocarbons.

4.3 Current databases and data handling practices

Whether the samples are obtained as part of the regular sampling effort of an observatory, or as part of a more sporadic cruise series, there are some conditions that should be fulfilled in order to maximize utility of the microbial deliverables gathered. There is a need for:

(i) Well organized repositories of microbial cultures and microbial genetic material, accessible to scientists with ideas on how to analyze the material;
(ii) Samples must be ecologically-referenced, with as much ancillary ecological and biogeochemical data (metadata) collected alongside with the microbial diversity data as possible; and
(iii) Sequence repositories which are geographically and ecologically-referenced.

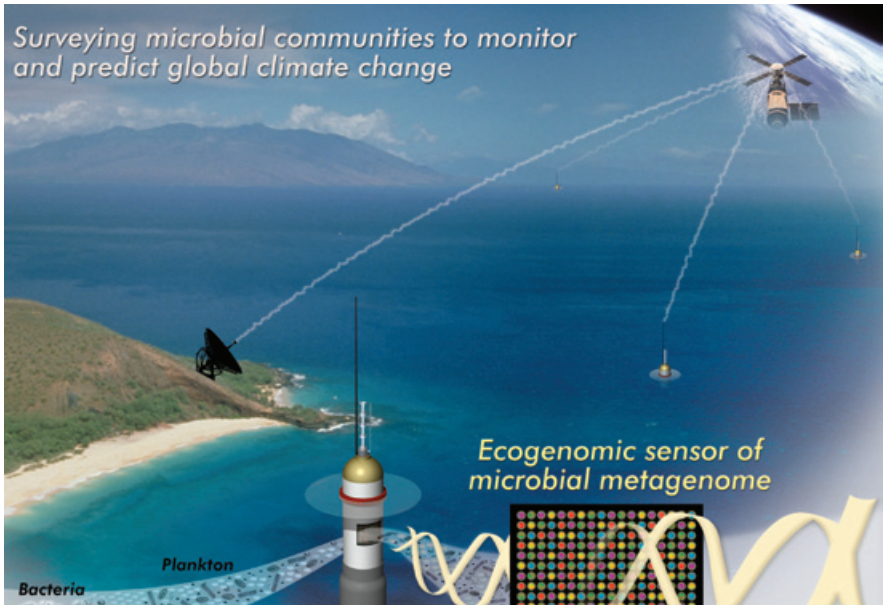


Figure 4.4. Illustration of an economic sensors embedded within an array of airborne and in situ sensors to be used for surveying microbial communities (courtesy U.S. Department of Energy Genomic Science program – http://genomicscience.energy.gov)

4.3.1 Reference culture storage
(#4 in Figure 4.1)

Classical marine microbiology still relies heavily on the culturing of microorganisms for identification and description of new species, and the biochemical characterisation of novel gene products. The isolated organisms are also useful as model organisms for ecological or physiological experiments. The cultures obtained during this type of research are valuable, but need constant care, and as the number of cultures grow, individual researchers cannot maintain these collections as part of their work. A solution needs to be found to store and maintain these culture collections.

This is currently provided by reference culture collections, such as the American Type Culture Collection (ATCC), the German Collection of Microorganisms and Cell Cultures (DSMZ), the Spanish Type Culture Collection, or the Institute Pasteur Collection[11]. These are institutions in which dedicated technicians take care of the samples of type strains and produce copies for the interested scientists. Nowadays, researchers that keep in-house culture collections might send materials to those who apply for samples depending on their availability of time, but they are not formally in charge of doing so. The type of repository that is needed is more similar to some algal culture collections (such as the Culture Collection Yerseke (CCY) at NIOZ or the Roscoff Culture Collection (RCC)[12]) that maintain a diverse array of ecologically relevant strains. Microdiversity and diversification studies require that a large number of similar microorganisms are being cultured and compared at the biogeochemical, ecological and phylogenetic level. This is not possible with standard culture collections.

Scientists should also be strict in that cultures and information about culturing requirements are sent to the repositories when a paper describing some aspect of an organism is published. Currently, we are faced with the paradox that some organisms have their full genome sequenced (and it can be downloaded from databases) but the organism has never been deposited in a culture collection in spite of the fact that several Nature papers were published using this non-deposited model organism.

What is needed is (i) central collections, such as specimen collections in museums with a rich set of electronically available metadata about the isolate; and (ii) a descriptive catalogue describing what is available and where. Both issues are currently targeted by the ESFRI project MIRRI[13].

11. http://www.lgcstandards-atcc.org/; http://www.dsmz.de/; http://www.cect.org/; http://www.pasteur.fr/ip/easysite/pasteur
12. http://www.sb-roscoff.fr/Phyto/RCC/index.php?option=com_frontpage&Itemid=1
13. http://www.mirri.org/

4.3.2 Genetic material repositories
(#5 in Figure 4.1)

General and well-organized repositories of microbial genetic material do not exist. They are at the home institutions of the responsible scientists, in various states, and there is no public access to them. However, the scientists in charge should be able to obtain financial support for these repositories from their institutions, their governments, or the EU. The rationale for this is firstly marine microbial genetic material is costly to obtain, requiring sampling on expensive research cruises. Secondly, repositories of historical samples could generate research by different groups or at different times, as techniques are continuously being developed and material collected years ago can be useful later on. We are convinced that the final model should promote repositories of cultures (culture collections) as cited above, in which a double funding model is used: some part of the financing is obtained from the users, but the basic expenses are paid by institution or government grants. The EU could use some of their research funds for the maintenance of these repositories.

An issue to discuss and study is the specific storage conditions in which the genetic material should be maintained. This is a tricky question as ongoing research tries to understand what makes DNA and RNA more or less stable and what types of fixatives can be used. Currently, samples tend to be stored in liquid N_2 or in -80°C freezers, but little is known about long-term DNA/RNA/protein stability under the most common storage protocols.

4.3.3 Sequence databases
(#6 in Figure 4.1)

The GenBank (NCBI) and EBI/ENA (EMBL)[14] sequence databases (International Nucleotide Sequence Database Collaboration (INSDC)) can be considered as data archives. Authors are obliged by journals to submit the sequences described in their papers to the sequence databases maintained by INSDC. Each sequence submitted receives an accession number, and these accession numbers are mandatory references in the publication. As the amount of information in INSDC increases with time, INSDC is less and less useful, because it contains increasing low quality (short) environmental sequences. Authors can then use the RDP10[15] database (which maintains a "classifier" function), or use the curated SILVA[16] database and Greengenes[17] repositories. These databases include only the fraction of rRNA gene sequences from INSDC which are of high quality

14. http://www.ncbi.nlm.nih.gov/genbank/ and http://www.ebi.ac.uk/embl/
15. http://rdp.cme.msu.edu/
16. http://www.arb-silva.de/
17. http://greengenes.lbl.gov/

according to the criteria set by the curators. A database which is specific for short-read sequences is the VAMPS database system.[18]

Metagenomic data are stored in specialized databases, such as CAMERA or MG-RAST[19]. MEGAN[20] is downloadable software that allows interrogation of the SEED and KEGG[21] genes databases. But it does not store the data. MEGX[22] is an ecologically oriented web portal that collects genome and metagenome data, displays it in a geo-referenced global system that combines the genetic data with ecological data obtained from the World Ocean Atlas and World Ocean Database[23]. This system integrates the SILVA database for rRNA gene-based diversity investigations.

Note that these systems are not a way for storing the data. Instead, they are portals to integrate, access, and compare the data.

4.3.4 Fingerprints and FISH/ARRAY samples (#7 and #8 in Figure 4.1)

There is no centralized organized way of maintaining fingerprinting images (e.g. DGGE gels) or data (e.g. ARISA profiles). Similarly, the maintenance of array data and filters to perform FISH is done in house in each laboratory.

4.3.5 Ancillary data (aka contextual or metadata) (#1, #2 and #3 in Figure 4.1)

Geographical and ecological accompanying data should also be stored using a standardized scheme. Basic data (geographical location, time) should always be accompanied by basic oceanographic data (temperature, salinity, chlorophyll or fluorescence). Nutrient concentrations, particularly inorganic nutrients, should be mandatory. Organic nutrients are also highly desired. Microbial abundances should be reported alongside the methodology used to obtain them.

Initiatives exist that can be used as role models: the Genomic Standards Consortium, an open-membership, international working body founded in 2005 provides mechanisms that standardize the description of genomes and the exchange and integration of genomic data. Marine data initiatives and databases (EMODNet, PANGAEA, SeaDataNet, Group on Earth Observations (GEOSS), Global Monitoring for Environment and Security GMES)[24] exist and new standards and mechanisms for seamless data transfer are being developed and implemented.

4.3.6 The flow of data

The diagram in Figure 4.5 sketches, in a simplified way, how data is generated from the successive research activities and projects in research institutes, how data are managed, and how these data are finally made available to the research community. In this scenario we will assume the research activities start with the samples taken, and measurements carried out on a research vessel during a scientific cruise.

4.3.6.1 Scientific cruises, projects and funding

Marine scientists need to reserve/book ship time on a scientific cruise organized by a ship operator. The operator will record this booking in an administrative database also called a cruise planning database**.** This information is not always public, mostly only available for the research operator, but some international initiatives share this kind of information. POGO and UNOLS share the cruise schedules of their research vessels in an attempt to optimize ship time. After each scientific cruise, but with some lag, the principal investigator of the cruise writes a cruise report describing what stations were visited, what kind of samples were taken, and what scientific work was carried out by the participants of the cruise/campaign. Some public databases have been built to capture this information. The ICES ROSCOP forms and the SEADATANET cruise summary reports capture the information from these cruise reports and make them available online. For research organized within a publicly funded project (national, European funds) the different progress reports will also show a trace of the activities undertaken, and the results so far. Sometimes administrative project information can be found in different data systems of science policy agencies and the EU funding administrations.

National marine data centres document these projects in a more scientific way, and some international networks have organized shared project information repositories. The global change master directory (GCMD) of NASA, and the European directory of marine environmental projects (EDMERP) started under the Sea-search project, and continued under SEADATANET to collect, standardize and make these project descriptions publicly available. This documenting activity generates what is recognized as “discovery metadata” allowing scientists to discover ongoing marine research projects, and contact their colleagues.

Publicly available cruise schedules, cruise summary reports, and project information is important to optimize

18. http://vamps.mbl.edu/
19. http://camera.calit2.net/ and http://metagenomics.anl.gov/
20. http://ab.inf.uni-tuebingen.de/software/megan/
21. http://www.theseed.org/wiki/Home_of_the_SEED and http://www.genome.jp/kegg/
22. http://www.megx.net
23. http://www.nodc.noaa.gov/OC5/WOA05/pr_woa05.html
24. http://www.emodnet-chemistry.eu/; http://www.pangaea.de/; http://www.seadatanet.org/; http://www.earthobservations.org/geoss.shtml;

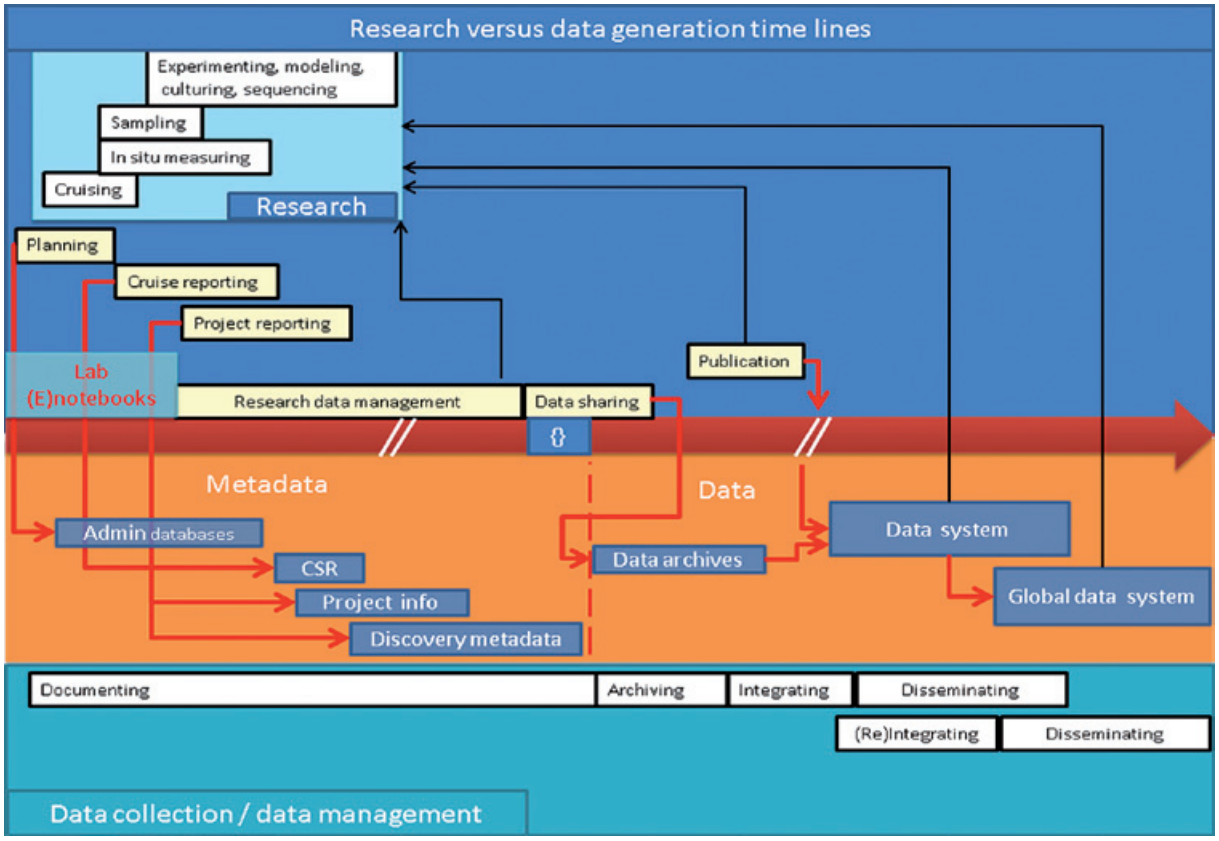


Figure 4.5. Data generation scheme providing a simplified overview of how data are generated from the successive research activities and projects in research institutes, how these data are managed and finally made available to the research community. The red arrow represents time. The dashes represent time gaps. The upper dark blue rectangle represents activities at the research institute, the lower blue rectangle activities at a data centre. The middle orange rectangle represents the different data products that are generated. (Developed by F. Hernandez)

expensive ship time and to enhance collaboration between scientists. The information collected is general, but is made available a long time before the actual data is released.

4.3.6.2 Data capture and documentation

The sampling history and analysis results are collected in laboratory notebooks (paper or electronic), often without clear guidelines. If the lab is better organized, these data are sometimes entered into a research database maintained by the most 'IT-minded' scientist in the lab or department. It is not accessible to the wider community, and unless it is documented by a data centre, its existence is not even known. Some national or international projects organize entire work packages to collect the data from the participating research groups. Unfortunately, the data collected in such projects do not always find their way to a professional data centre, and risk getting lost after the project ends. In data management terminology this collection of data is known as a 'dataset', and the description of a dataset is called 'metadata'. (Depending on the rules used to group data, a dataset can be everything collected during a thesis, a project, a specific campaign, or even all the data collected by a research group during several decades).

Data centres describe datasets according to several standards (ISO19115, Dublin core) and at different levels of detail. The most general level is called 'discovery metadata' and its purpose is to make the dataset 'searchable', 'discoverable'. The metadata describes the content of the dataset (which parameters it includes), the period and region where it was collected, who collected the data, where it is stored and under what conditions it can be shared (a summary of the data policy).

Some international metadata directories include: The register of resources started under the MARBEF NoE, to describe all known marine biological data sets and forms the basis for EUROBIS; the European Directory of marine environmental data (EDMED) which was initi-

ated by BODC in 1991 under the EU MAST framework, and collects discovery metadata of European datasets described by the National Oceanographic Data Centres in Europe, and is used by SEADATANET. The global change master directory (GCMD) maintained by NASA is a global environmental (meta)database.

4.3.6.3 Publications, data and information databases

Many scientific results described in publications never find their way to data centres or archives. And in many cases, even if the publication is electronically available, the information within is not sufficiently structured to be re-used. This is mainly due to the fact that no common rules exist for electronic data publishing. Consequently the user is faced by a mixture of different data formats, access policies and submission procedures if there is an appropriate database at all.

When species are described in publications, a plethora of methods for naming, classifying and referencing taxonomic entities has been used, but until recently, no structured lists of species names were available, which resulted in numerous multiple first descriptions of the same species. Taxonomic databases, like ALGAEBASE maintained at the University of Galway, the world register of marine species (WORMS) maintained by VLIZ, the catalogue of life (CoL) maintained by SPECIES2000, the pan European species infrastructure (PESI), an FP7 project, are far from complete in terms of the number of species names, let alone in terms of species information.

Particularly for the microbial taxa, databases are incomplete, which is partially due to the lack of a practical "species concept" for microorganisms. The rules for the formal description of a new microorganism in the literature require culturing and biochemical characterization of the strain, which is not only time consuming, but often technically impossible. The commonly used species registers lack the possibility to deal with non-taxonomic entities like operational taxonomic units (OTU), which poses problems in the standardization steps needed for data integration of sequence and meta-omics observation data collected by microbiologists.

4.3.6.4 Archiving Oceanographic datasets

The purpose of data archiving is to make sure the data is preserved for future use. Scientific data sets will eventually get lost if not properly archived, because the originating scientist left the institute, or because the media on which it was stored becomes unreadable. Data archiving implies that the data set is documented, and stored in a safe location. The data are often stored as files, without need for further standardization, though some data centres have developed ways of capturing this heterogenic data in structured databases.

National Data Oceanographic Centres within the IODE networks are mandated to collect and archive datasets. The European data centres have organized themselves in a real network thanks to the Sea-search project, and subsequent SEADATANET projects.

The publishing network for geo-scientific and environmental data (PANGAEA), is hosted by AWI and MARUM, and used by the world data centre for marine environmental sciences (WDC-MARE). It is one of the biggest European marine data archives. Besides dealing with project data, it also developed a publishing system for archived data using references based on DOI's.

Several years can pass before the data flows from the researcher to a public data centre. This is often dictated by the project funder, sometimes by the publisher of a journal. The National Science Foundation requires that any funded project entrusts the generated data to a permanent data archive. The datasets in a data archive are not publically accessible immediately. Access can be restricted to a research group, or project partners. Public archives also offer a moratorium period, during which the submitted dataset is not publicly available. This period is normally linked to the date of publication. The dataset is then made available for download from the data archive, or from the integrated data system.

4.4 From data to integration and interpretation

The next stage in data management involves several actions (standardization, quality control, biocuration) to combine datasets from different sources (data providers) in one data system. This succession of activities is also known as 'integration of datasets'. This integration gives higher value to the individual datasets; higher values depending on the actions taken for each specific data system.

The data are at least transformed to a common format and quality controlled. Standardization of the parameters measured, the units used, the time and location and depth of the observation, the taxonomic entity, the chemical compound are also essential for combining data from different sources.

The deliverables collected during field work are finally analyzed and interpreted and the final product (new knowledge, whether published in manuscripts or not) is created from the genomic (*sensu lato*) and environmental data, originating from field studies, lab analyses or after computer work.

Again, the linkage between the final product and the original data is often lost or, at least, cannot easily be

traced. Raw data, except for sequences deposited in INSDC or sites like CAMERA, are very often absent from the published manuscripts, and while at times raw data are deposited in the Institutes' websites, with time they are likely to be lost.

4.5 Concluding remarks

Microbial biodiversity data obtained from samples of opportunity or from samples taken on multipurpose oceanographic cruises are too often collected in a non–standardized way, less than optimally documented, and fail to capture the temporal variation in microbial communities or their relation to environmental parameters. Scientists research agendas and goals should take this into account and adapt accordingly. It is not that the current studies are not valid, but using organized sampling systems would add considerable value to them. A few observatories collect time-series data from fixed stations, but their monitoring programs often lack microbial diversity data, and, when they do, their approaches are not coordinated. For many research applications the environmental (geographic, oceanographic, biogeochemical, biodiversity, experimental manipulation) data need to be collected at the same spot and time of the microbial sample. In such cases, multidisciplinary sampling is essential.

Summary Box 4.1. Key research priorities and recommendations

1. Develop a Pan EU long-term programme for multidisciplinary sampling of each European sea for at least 10 years. This should cover at least one microbial observatory from each of the European coastal and deep oceans and also deep-sea microbial observatories.
2. Stimulate the development of durable partnerships between the marine microbial research community and industry to develop and apply novel data acquisition technologies involving microbial diversity and function data. Automatic data acquisition technologies will become important tools in the future. It should be the aim of the European scientific and industrial partners to lead the development and application of this innovative technology in the future.
3. Develop new standards and ontologies, and develop or enhance existing data systems in order to capture and interlink marine microbial information with environment, habitat, genomic and species information registers. New models need to be developed to generate the missing data products (derived data).
4. Enhance existing data and metadata repositories to better fulfil the requirements for marine microbiological applications. Many sequence data entries available in public data repositories (INSDC) lack the proper geo-referencing and environmental annotation needed for ecological studies.
5. Increase efforts to raise awareness and develop strategies for the scientific marine microbiology community to improve the (i) collection and reporting of ancillary data, (ii) storage of DNA, cultures, microbial biomass, etc. in a well-organized repository, (iii) linking samples to the ancillary (geographic, ecological, oceanographic, experimental) data, and (iv) respect for international standards for data reporting, and assure the maximal linkage between all products of research (i.e. papers, cultures, data, stored DNA, sequences, etc.). Training efforts should be directed at the PhD level.
6. Organise repositories for all samples, cultures and genetic materials collected during a project, and find permanent facilities to host them. Samples, cultures and genetic materials are valuable as reference for future research, but need constant care and the individual scientists cannot maintain them as part of their work.
7. Funding agencies need to contribute to address the loss of data collected by public funds and specify requirements for each project to permanently archive and publish data generated by the project in public data systems. Much of marine biological, microbiological and environmental data are still lost because they are stored on perishable media, or buried in local archives which never become public.
8. Build and run permanent (long-term funded) dedicated data infrastructure nodes for microbial marine data analysis linked to existing data initiatives like EBI, EMODNET, WORMS, LIFEWATCH. With the explosion of Omics data, we will need faster, more automatable data exchange systems. The data delivery systems based on 'discover, request copy and download' are fine to provide access to some datasets, but are not suitable for the machine to machine interfaces that are needed to automate the data processing and analysis of these huge data streams.

5. European status, research gaps and recommendations

5.1 Introduction

The importance of marine microorganisms, their diversity and impact, as well as advances in research toolkits and infrastructures have all been addressed in this report. Previous chapters have revealed important progress in our ability to address the societal and research questions introduced in chapter 1 (see Information Box 2). Important societal questions include how marine microbial research can contribute to addressing some of the most important challenges of our time such as human and environmental health, global change, sustainable supply of healthy food and energy and supporting a thriving European bioeconomy. However, major gaps remain in our understanding of both the diversity and functioning of marine microorganisms which hampers our ability to exploit the full range of opportunities that marine microbial research can provide for our societies. Research and technology development are the key to unlocking this societal potential. Before making recommendations for future research priorities and enabling actions, it is necessary to address the following key scientific questions:

1. What is the nature of microbial diversity in the ocean and seas?
2. Which taxa are most relevant in terms of function across different ocean ecosystems?
3. Which metabolic pathways exist?
4. Which regulatory and signalling networks exist within and between microbial communities?
5. How do environmental factors influence metabolism and regulation within and between taxa?
6. How do microbial interactions influence ecology and ecosystem functioning?

Because these scientific questions are central to our ability to gain a deeper understanding of marine microbial diversity and to reap the societal benefits associated with that knowledge, this chapter uses them to structure a review of the European marine research status, existing gaps and recommendations for future marine research programmes.

Developments in the area of marine microbiological research in Europe, and by extension marine research in general, cannot be addressed in isolation from the institutional, funding and science policy landscape which has evolved in the last decades, both in Europe and beyond. Therefore, this chapter begins with a reflection on European marine microbial research with respect to the broader European science policy context and global developments.

5.2 The status of European marine microbial research

5.2.1 Science policy context

As pointed out in the Galway Declaration (2004) and reinforced by the Aberdeen (2007) and Ostend (2010) Declarations, marine science and technology plays an essential role in generating the knowledge needed to support the prosperity and well-being of Europeans, especially in times of global economic, energy and environmental crisis.

Influenced by these initiatives and advances in science and technology, the marine science policy landscape has developed significantly during the last 10 years. The Maritime Green Paper (2006) which developed the vision for the Integrated Maritime Policy (IMP, 2007) highlighted that marine research, and in particular marine biotechnological research, is a driver of economic activity as it generates new knowledge in many scientific and technological disciplines. Therefore, the impact of the outputs extends much further than the marine field: it stimulates activity in areas such as health, food, energy, pharmaceuticals, environment and transport.

In September 2008 the EC delivered a European Strategy for Marine and Maritime Research as the research pillar of the European Integrated Maritime Policy (IMP). It again prioritized marine biodiversity and biotechnology research and recognised their potential to contribute new knowledge, on which to base high value-added products and processes supported by excellent scientific research.

Marine microbiology research has grown considerably during the past 15 years, partly as a result of the wider recognition of the enormous potential of marine biotechnology for the development of new applications and products. This has been reflected by many science policy initiatives and events. For example, as a result of the Bremen Meeting on Marine Biotechnology Research (2007), the European Commission set-up a Collaborative Working Group on Marine Biotechnology (CWG-MB) which highlighted the potential of the sector and provided clear recommendations to support its development in Europe. The EC-CWG culminated in the FP7 Coordination and Support Action (CSA) MARINEBIOTECH (www.marinebiotech.eu) which is preparing (2011-2012) for an ERA-NET in the field of marine biotechnology research.

In 2010, the Marine Board organised an ESF-COST High-Level Conference on Marine Biotechnology and published a position paper resulting from the work of an expert working group on marine biotechnology. In

2008 and 2010, EC-US Task Force on Biotechnology Research workshops in Monaco and Washington respectively, emphasized that microbes play a critical role in the marine and global ecosystem and are considered responsible for a significant proportion of biological activity. In addition to highlighting research challenges and opportunities, the workshops also stressed the need to strengthen Europe's bioinformatics research capability.

In the last decade, the European Commission has recognised the need for European coordination and collaboration in marine microbial and biotechnological research and has funded, through its Framework Programmes, several networks and projects designed to support and stimulate research on marine diversity and genomics. Two Networks of Excellence (NoEs), Marine Biodiversity and Ecosystem Functioning (MarBef) and Marine Genomics Europe (MGE), as well as a range of integrated research projects such as HERMES (Hotspot Ecosystems Research on the Margins of European Seas) were able to bring together more than 150 institutions, over 1000 researchers in over 30 countries (within and outside Europe). Also at the pan-European level, bringing together funding from national research funding agencies, the European Science Foundation (ESF) managed a multi-national European investment in the EuroDEEP initiative which aimed to describe, explain and predict variations of biodiversity within and between deep-sea habitats, with respect to deep-sea ecosystem functioning and the global biosphere.

By funding the ASSEMBLE (2009-) research infrastructure initiative and the EUROFLEETS (2009-) infrastructure project, the integrated projects MAMBA (Marine Metagenomics for new Biotechnological Applications, 2009-), EuroMarine (2010-), MicroB3 (2012-) and MaCuMBA (2012-), the EC has taken action to specifically support marine biology and biotechnology in the 7th framework programme. In particular MicroB3 (2012-) and MaCuMBA (2012-) are specifically addressing challenges associated with marine microbial research. These major initiatives and several other European projects have contributed to the initial development of a European Research Area (ERA) for marine sciences. The concept of the ERA, first coined at the Lisbon European Council in March 2000 and given new impetus in 2007 with the European Commission's Green Paper on ERA, aims to (i) improve dynamism and innovation in all sectors of industry and services, resulting in more and better jobs; (ii) address important issues of a European and even global dimension, such as health, energy supply and climate change; and (iii) create a society in which knowledge is shared, taught and valued as an essential source of personal and collective development.

Interdisciplinary approaches, ranging from sampling and lab work to bioinformatics is a prerequisite to be able to address the complexity of the marine system. Integrating marine sciences and bringing it from an institutional to a European level has significantly improved the networking between marine researchers and institutes. The policies and programmes which have successfully been implemented over the last years have laid a cornerstone in our capacity meet the challenges of the 21st century in marine biology. The Tara and Malaspina cruises (see Information Box 7) can be seen as a first spin-off of this new kind of thinking in marine sciences.

5.2.2 Global developments in marine microbial research

Globally, marine microbial research has gained significantly more attention in recent years. One of the most prominent projects of the last decade was the Global Ocean Sampling Expedition (GOS) by J. Craig Venter, an ocean exploration genome project with the goal of assessing the genetic diversity in marine microbial communities and to understand their role in nature's fundamental processes. The project started as a Sargasso Sea pilot sampling project in August 2003 and in 2004 extended to a two-year global circumnavigation using Craig Venter's personal yacht, the Sorcerer II. The first part of GOS, which sampled the North Atlantic, Caribbean and a small part of the Pacific Ocean, added DNA sequence information equivalent to 50% of all protein-encoding sequences that had previously been deposited in the INSDC (International Nucleotide Sequence Database Collaboration). GOS confirmed that marine microbes are diverse, revealing how little is known about the genetic information of natural assemblages.

In terms of sequencing capacities, Europe currently lags well behind the US and Asia. The Joint Genome Institute (http://www.jgi.doe.gov/) as well as the Broad institute (http://www.broadinstitute.org/) in the US are the most prominent in the delivery of large scale environmental sequencing projects and since 2008 the BGI (formerly Beijing Genomics Institute, http://www.genomics.cn) offers a sequencing service for nearly anything that contains DNA. In Europe only two main sequencing centres are available: Genoscope in France (http://www.genoscope.cns.fr) and the Sanger institute in the UK (http://www.sanger.ac.uk/). With Sanger exclusively focusing on medical applications, Genoscope is the only resource for mid- to large-scale environmental sequencing approaches.

The requirements for data management and analysis are increasing rapidly, mainly as a result of the growing and more diverse application of omics technologies.

Information Box 7. European circumnavigation cruises

Stimulated by the success of the GOS and Sorcerer cruises, Europe recently launched two exciting circumnavigations with the Tara- and Malaspina cruises. Tara-Oceans is a three year (2009-2012) circumglobal expedition being carried out on the 36 metre research schooner Tara (www.oceans.taraexpeditions.org). The project is using satellite-derived information to target mesoscale and sub-mesoscale structures of scientific interest throughout the Earth's ocean. The scientists involved in the project have established a standardized sampling protocol designed to sample plankton communities from end-to-end, including viruses, *Bacteria*, *Archaea*, protists, and zooplankton. The project's aim is the reconstruction of ecosystem community composition and understanding its functional potential within the physico-chemical constraints of each sampling site.

The Malaspina circumnavigation expedition is an interdisciplinary research project designed to assess the impact of global change on the oceans and explore their biodiversity (http://en.wikipedia.org/wiki/Malaspina_Expedition_2010). The 250 scientists on board the Hespérides and Sarmiento de Gamboa have embarked on a nine-month expedition (December 2010 to July 2011) combining pioneering scientific research with training for young researchers, while advancing marine science and fostering the public understanding of science. The project is under the umbrella of the Spanish Ministry of Science and Innovation's Consolider – Ingenio 2010 program and is led by the Spanish National Research Council (CSIC) with the support of the Spanish Navy. Malaspina's main objectives are: (i) to assess the impact of global change on the oceans, (ii) explore the biodiversity of the deep ocean, (iii) assess the impact of the original Malaspina expedition, (iv) promote marine science in Spain and public understanding of issues in marine sciences and (v) to raise the interest for marine sciences among the youth and training of young scientists in a global perspective to ocean sciences.

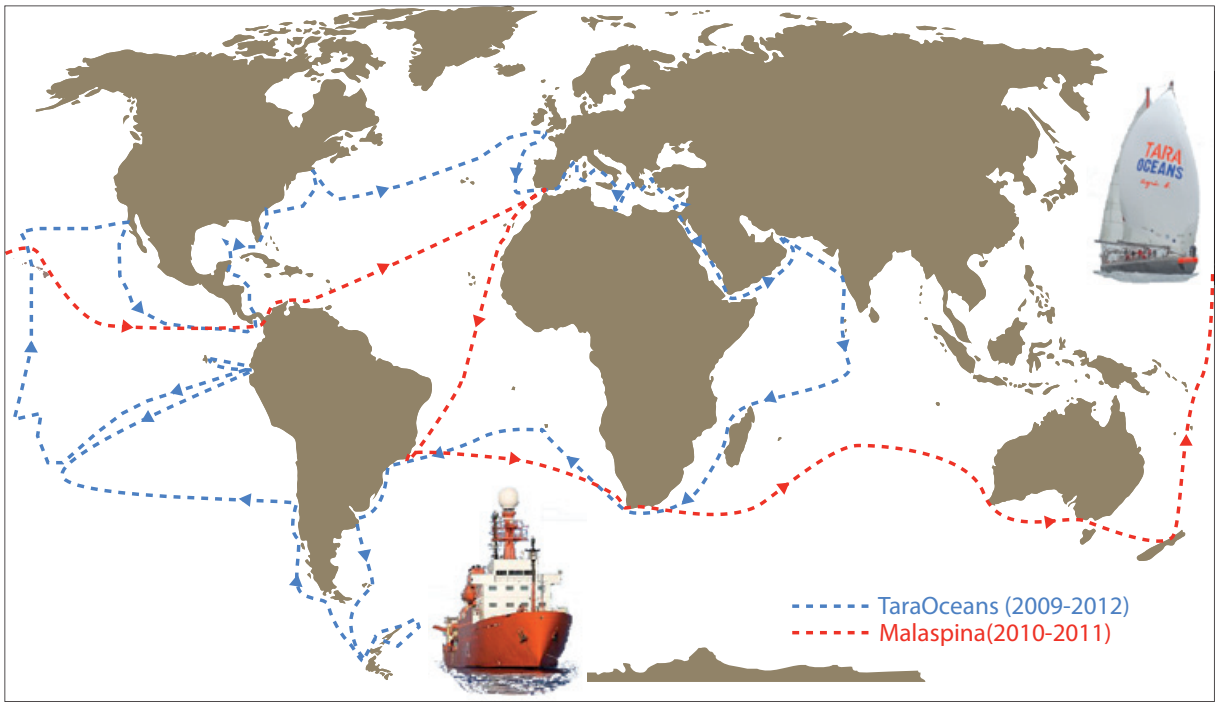


Figure 5.1. Overview of the TaraOceans and Malaspina circumnavigation cruises (image by B. Garriz, E. Broglio and J.M. Gasol)

These technologies will result in a huge increase in the amount of sequence and environmental data in the near future which will need to be properly managed to allow analysis and meaningful research results. Currently, the GOS circumnavigation data represents the largest metagenomic dataset ever placed in the public domain, with more than 13 million sequences and over 12 billion base pairs of DNA. It is clear that the data of just the European cruises can easily extend into the Petabyte range (i.e. 1,000,000,000,000,000 bytes).

To cope with the size and complexity of such datasets the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) has been created in the US with funding from the Gordon and Betty Moore Foundation. Also in the US, the Integrated Microbial (Meta)genomes (IMG/M) and MG-RAST (Metagenome Rapid Annotation using Subsystem Technology) platforms hosted by the Joint Genome Institute and the Argonne National Lab have proven extremely useful for data analysis and annotations, although not specifically created for the marine sector.

In Europe, only EBI-ENA (EMBL) has the capacity to store and make available at least parts of the data. For bioinformatic data analysis, no dedicated marine bioinformatics infrastructure like CAMERA exists in Europe and data analysis is mostly carried out by the biologists themselves. This situation has advantages in terms of flexibility and the intimate connection between the environment, samples, analysis and interpretation. Unfortunately it also leads to heterogeneous approaches and "island solutions" which hamper large-scale data integration and interpretation. Furthermore, it is often difficult for individual research groups or even marine institutes to reach a critical mass in bioinformatics that can allow them to cope with the flood of data produced by omics technologies.

5.3 Major research gaps and challenges for future research

Although unprecedented technological improvements in the last years have provided us with a wealth of knowledge about the dominant types of microorganisms in the oceans and their activities, there are still many research questions that have not been solved, either because the appropriate methodologies have not yet been developed or applied or, more often, because answering these questions implies an amount of work and resources that is out of the reach of most laboratories.

The following key research priorities and needs are proposed with respect to the main high-level **scientific questions**:

(1) What is the nature of microbial diversity in the oceans and seas?

Marine microbial biologists are often confronted with the question: "How many microbial "species" exist in a given environment?" In microbiology the real biodiversity can, at best, only be estimated using an operational approach to approximate a microbial "species". This is usually done by cultivation-independent techniques, mostly based on the analysis of the ribosomal RNA genes. Classical clone or single-cell based approaches cannot give a comprehensive picture of microbial diversity due to the low statistical significance of the results. Nevertheless, first usage of next generation sequencing (NGS) technologies suggests that any given ecosystem is harbouring on the order of 10^4 microbial "species", which means that most of the species show only very low abundances and comprise the so called "rare biosphere". This concept has important ecological implications, as it would indicate that most bacterial types exist everywhere and that most ecosystems have the potential for responding to changes in environmental conditions. For example, in the case of an oil spill, the rare biosphere may contain organisms adapted to the use of hydrocarbons which can become more abundant. Recent reconsideration of the errors involved in pyrosequencing have decreased the initial estimates of how many different "species" an ecosystem could have, but the numbers are still impressive and in the order of thousands per sample.

The range of uncertainty expressed above indicates the clear need for increased research in marine biodiversity. Next Generation Sequencing now allows screening for biodiversity for any "marker gene" at unprecedented temporal and spatial scales. This could be used as the starting point to determine the dynamics of microbial communities as well as detecting recurring patterns of microbial association and potential interactions. This might provide an answer to the decades-old question of whether everything is really everywhere. After a baseline of biodiversity in the oceans has been determined, anthropogenic influences that change biodiversity could be distinguished from natural fluctuations. This would also allow for monitoring the spread of potentially pathogenic bacteria triggered by warming of the ocean surface.

(2) Which taxa are most relevant in terms of function across different ocean ecosystems?

Recent research indicates that marine systems harbour groups of organisms that are distinct from the ones found in freshwater and terrestrial habitats. At least one group, the alpha subclass of *Proteobacteria,* shows a preference for the marine environment. The SAR11

clade, which can be found in nearly every marine sample, is a prominent example of this group. Although a rather broad clade, even *Euryarchaeota* have turned out to be a ubiquitous group preferentially found in marine systems.

Although there are clear indications that core taxa exist, there is also evidence that changing environmental conditions with water depth provide a stronger vertical than horizontal determinant of microbial biodiversity and distribution and that certan microbes are confined to certain water masses. However, it appears that microbes are seasonal and that each year some appear in winter while others occur in summer. What drives this succession? What are the differences in the microbial genomes that make a given SAR11 representative be a winter "species", while another one prefers the spring temperatures?

We also need to understand how boundaries in the ocean are formed which contain microbial communities of similar structure. Like terrestrial ecologists we should be able to define the biome and characterize the ecological traits of the species that thrive in each biome. What are the boundaries of the "ocean biomes"? This can only be answered by a combined extensive effort with an intensive effort at many sites: the microbial observatories. Several such observatories are already operating in Europe including the L4 Plymouth site, the Roscoff station, Helgoland Roads in northern Europe, and the Blanes Bay, Mola and Sola and Dyfamed stations in the NW Mediterranean. However, this does not yet represent a sufficiently comprehensive or coordinated network and the maintenance of these sampling sites currently relies too heavily on national funding which makes them more exposed to the economic crisis.

The current lack of detailed data on "ocean biomes" hampers our ability to determine which functions and taxa are most relevant in the different ecosystems. With the Earth Microbiome Project (EMP) the Argonne National Lab in the US has recently begun to investigate the diversity and function of the microbial world more systematically and from the viewpoint of a microbe (www.earthmicrobiome.org). EMP plans to sequence 200,000 samples resulting in a total of 2.4 petabases of sequence data comprising of DNA, mRNA and rRNA. While only a subset will be dedicated to the marine system, this could nonetheless deliver first indications of whether "ocean biomes" exist. An important advantage of the EMP project is that sample preparation and acquisition of contextual (meta)data will be centralized and standardized, so that a direct comparison of the results will be possible.

(3) What metabolic pathways, regulatory and signalling networks exist?

Omics technologies open the possibility to provide an inventory of enzymes, pathways and short regulatory nucleotides of the marine system. With metatranscriptomics and metaproteomics it is even possible to get insights into the activity of genes at a certain point in space or time. Nevertheless, because the reconstruction of enzymatic functions and metabolic pathways using omics technologies is mostly based on bioinformatic predictions, the determination of the exact specificity of the enzymes is still impossible. Additionally, in each genome or metagenome investigated so far, up to 50% of the predicted genes could not be assigned with a function. These hypothetical or conserved hypothetical genes most likely harbour the information of the specific adaptations of the respective organism or community. Single cell technologies can help to assign these genes to specific taxa, but to elucidate their functions more cultivates are a prerequisite.

If the quantity and quality of data increases substantially, it may be possible to investigate the extent and spreading of microbe-driven ecological functions in microbial genomes. This relates to the resilience of a microbial community to external forcings. Are functions duplicated or multiplicated, or are there only a few organisms performing each function? Research to date suggests that this depends on which function is considered: organic carbon respiration is probably widespread and ecosystems are resilient in terms of performing this function, while it is likely that the utilization of a specific wavelength for phototrophic growth, for example, is less common.

A dense network of environmental and genomic data could also provide clues about the functions of hypothetical genes by correlating environmental "biome" features with recurring gene patterns of unknown functions. Such "guilt by association" inferences are already known to e.g. determine potential functions for genes as part of transcriptional units (operons), but the approach of using environmental parameters to predict functions is in its infancy.

(4) How do environmental factors influence metabolism and regulation within and between taxa? How do microbial interactions influence ecology and ecosystem functioning?

The overwhelming and fascinating complexity uncovered by omics techniques has become a major focus in contemporary marine microbial ecology. The obvious challenge to marine microbial ecologists is to transform this avalanche of data at the molecular scale into "knowledge and understanding" about the complex

interactions of marine organisms, their metabolic and regulatory mechanisms, and ecosystem functioning in general.

Following a traditional hypothetical-deductive perspective, science can be seen as a spiralling process whereby observations lead to models (insights, hypotheses, etc.) which can be used to construct new experiments that generate new data, and where each stage brings us a further step towards better understanding. Looking back at the last decades of marine microbial ecology, however, an important refinement to this picture becomes obvious: journeys into new and unknown territories require open-minded (i.e. model-independent) observations to allow an unbiased observer to be able to see the unexpected and thus create a new entrance into the hypothetical-deductive spiral. For example, you need to know that there are ~10^9 bacteria and ~10^{10} virus per litre of seawater before you can start asking for mechanisms behind and relationships between such data.

A characteristic trait of a mature field of science is a common theoretical framework within which observations can be organized and related, new observations predicted, and experiments thus planned to challenge the theory. Despite its progress, one can ask whether marine microbiology scores particularly well according to such a criterion. To a large extent the field is probably still in the initial descriptive phase, gathering increasingly detailed descriptions of the unknown. Significant efforts are being invested to measure microbial diversity, but to what extent is this still the initial hunt for a correct number, more than a hypothesis-driven search for the underlying mechanisms and relationships producing a particular diversity? We know quite well the abundance of viruses, but do we even have a satisfactory theory explaining the order of magnitude of this number? We can do meta-analyses generating correlation plots between bacterial production and primary production, but to what extent can we match these correlations with mechanistic explanations for how the system works? Is the correct description of the state-of-the-art that we have a lot of observations and data on how the marine microbial ecosystem looks in terms of numbers, fluxes and even species, strains and genes, but still a relatively vague idea of the relationships between these numbers? From this perspective, the new wave of molecular data is flooding into a field that is immature in the sense of not having an advanced theoretical framework within which these data can be arranged, analyzed and interpreted.

In short, even with the vast amounts of data already available, our hypothetical approaches are not sufficiently advanced to permit large-scale predictions about the influence of environmental factors on the diversity and function of microbes and the influence of microbes on ecosystem functioning. However, when we lower our expectations from a "universal model" towards the more practical construction of tools with predictive power for applied use in e.g. coastal zone management or climate change prediction, a detailed knowledge about the microbial biodiversity and its functions in conjunction with environmental parameters and anthropogenic forcing is becoming possible.

Summary Box 5.1.
Strategic research priorities and needs

1. Ensure appropriate and accurate data acquisition to obtain more and better data;
2. Build a registry for samples and genetic materials;
3. Develop innovative cultivation approaches and, in turn, establish model organism databases and genetic systems for marine organisms;
4. Improve classical methods/approaches, develop novel techniques and combine their results with abiotic and biotic information. By applying such an approach it will become possible to generate a more precise picture of the function, interaction and diversity within marine microbial food-webs;
5. Generate a dense network of data on the marine ecosystem by increasing both the number and the technological capacities of marine observatories;
6. Contextualise sequence data with a set of standardized parameters (metadata) of habitats
7. Improve data management to be able to coordinate sampling, sample analyses and processing of data and make this data openly available for users;
8. Ensure interoperability of multidisciplinary data repositories;
9. Implement well-curated and integrated ecological/diversity/and genetic databases;
10. Fully implement a systems approach to microbial ecology to understand ecosystem functioning;
11. Prepare European researchers to be able to participate in, and benefit from, data delivered by international projects and programmes.

5.4 Recommendations

Recommendation 1.
Establish a coordinated pan-European research programme on marine microbiology

To address the societal and research questions raised for marine microbial research on biodiversity and ecosystem functioning at a comprehensive scale, a long-term, pan-European research programme is urgently needed. Now, at the onset of the data deluge, this would prevent fragmentation and redundancies in sampling and analysis, promote unified standards and enhance interoperability and data integration. Advanced research efforts must cross boundaries and involve a range of disciplines including, for example, chemistry, oceanography, ecology, genomics, information technology as well as social sciences and even legal experts to address issues related to intellectual property rights. Such a large-scale programme should not displace the funding of excellent individual investigators which has been central to many scientific advances in marine microbiology, and to science in general. It should rather be taken as an opportunity to promote marine microbiology and embrace the opportunity to recruit and train the very best students for a stimulating career in marine microbiology research in line with Recommendation 1 above.

The proposed programme should:

- Have a duration of 10 years;
- Involve regular in-depth sampling and analysis at up to 10 sampling sites selected at the onset of the programme. These could be part of one major project or a coordinated cluster of smaller sub-projects. The selection should follow a screening procedure to find a representative set of natural “ocean biomes”. Some of the sampling sites should have a history of data collection with respect to biodiversity, microbial and environmental parameters;
- Sample selected sites at least once per week to monitor changes in biodiversity and, at times of identified fluctuation, apply the full range of omics technologies;
- Develop a set of standard operating procedures for sampling, contextual (meta)data acquisition, data management and processing;
- All data should be made publicly available in a “one-stop-shop”, using the EBI-ELIXIR infrastructure, as soon as they appear in a solid state;
- From the outset, biological data should be backed up with oceanographic data from profile and remote sensing systems. Moreover, automatic sampling and data processing procedures should be developed in parallel, integrating relevant industrial partners.

As an overarching task, specialized and reliable databases for marker genes as well as functional genes should be established and maintained by a set of highly qualified bio-curators. The databases would help to establish a clear baseline of biodiversity at the sampling stations based on the analysis of marker genes. Specialized databases for e.g. commercially useful enzymes (food, energy, medicine) as well as specific metabolic, functional or resistance genes could be useful both in providing opportunities to develop commercial products and processes and in monitoring the status of the environment in times of global change.

The programme should also involve appropriate legal and policy expertise to address issues concerning access and benefit sharing, intellectual property, and data as well as for the exploitation of high-potential commercial applications.

The outcome of such a programme is expected to be the following:

- A set of community-agreed standards and technologies for (automatic) sampling and data acquisition (lab protocols), storage and exchange of data to reach a new level of interoperability and data integration across disciplines;
- Innovative software approaches for quality management, data processing, data integration, accessibility and visualisation;
- Ecosystem models for selected sites in the marine system to provide a predictive understanding of the contributions of functional microbial biodiversity to marine ecosystems functioning, with a special focus on the role of microbes in climate change and the effect of climate change on microbial communities;
- Delivery of new biocatalytic processes, enzymes, biosynthetic pathways and bioactive compounds for use in biotechnological applications.
- An innovative legal framework and model contracts for the protection and sustainable use of marine genetic resources;
- A series of workshops, outreach and training initiatives to make the project results accessible for researchers, industry, policy makers and the public at large;
- A new generation of marine microbial researchers well trained in oceanography, data management, omics technologies and biodiversity policy;
- A new communication culture with bilateral understanding between computer scientists, biologists, biochemists, bioinformaticians, oceanographers as well as philosophy and legal experts.

The integrated programme will not only provide answers to the high level societal questions and ecosystem management; its interdisciplinary character will cross

boundaries and support the development of a new culture of science communication. A better understanding of the problems and needs of the different disciplines and attitudes represented in a pan-European programme will help in preventing fragmentation in the European marine research area by substantially increasing technological and knowledge levels.

Recommendation 2.
Create a European repository for cultivated collections and a reference library

Even when we are capable of obtaining the full genome sequence of uncultivated microorganisms through next generation high-throughput metagenomic sequencing, and can monitor their activities through transcriptomics, proteomics and metabolomics, and have the techniques to visualize these uncultivated microorganisms in their natural environment, we will realize that "cultivation-independent" strategies will only answer part of our questions. To fully understand ecology and life-cycle of a microorganism and its full biochemical and acclimation potential, isolation and cultivation are essential.

Unfortunately, there are currently no requirements for scientists who publish results using not yet available microorganisms to make these microorganisms available to the scientific community. In order to reproduce the published results or to build on those results and continue research, scientists depend on the willingness of the owner of the strains to make them available which can create problems. For example, the owner may plan to continue research on that organism, or the strain has been lost, or the owner is simply unwilling to cooperate or respond.

Furthermore, the isolation, purification and cultivation of microorganisms in the laboratory is often very difficult, tedious and laborious. Too often, at the end of a project, the laboratory does not have the capacity or willingness to maintain such cultures and they are lost. Moreover, existing culture collections are often only interested in taking a small selection of isolates, such as novel microorganisms with unusual properties or of proven biotechnological interest, because they are limited in their financial and personnel capacities. As a result, several laboratories have started their own culture collections, often specializing in specific groups of organisms. Where they exist, these collections are frequently financed by short-term research funds and such collections usually disappear when the scientist in charge leaves or retires. More importantly, these culture collections usually lack the capacity to supply cultures to the broader scientific community.

Hence the following key steps are recommended:

- Create a virtual European repository and reference library for cultivated marine microorganisms with a single on-line portal and up-to-date catalogue (containing all the information of the strains, their origin, their growth media and cultivation conditions, (partial) genome sequences, images, etc.) and access procedure. Such a virtual repository could provide information and access to all significant European marine microbial culture collections and would require long-term financial support in cooperation with the home institutes. The virtual repository will ensure the quality control of the collections and control the prompt execution of orders for strains by clients;
- Strengthen local culture collections to provide expert knowledge and technology to:
  - Investigate the taxonomic position of the individual strains in the collection;
  - Develop growth media and optimize cultivation conditions;
  - Develop long-term storage protocols for all strains (e.g. by cryo-preservation or lyophilization) in order to minimize the labour required for maintaining growing cultures;
  - Maintain a DNA library of the collection to supply research groups who only need DNA of a particular strain.
- Provide training for culture collection curators and others involved in maintaining cultures of microorganisms as well as for young scientists and research assistants in order to educate the next generation of dedicated culture collection curators.

Recommendation 3.
Create an integrated, multidisciplinary European Centre for marine data management and analysis

To address the research questions 1-4 (see Information Box 2, page 15) and to prevent that the large-scale implementation of omics technologies becomes disruptive (where the data production is faster than the speed users are able to make use of it) it is recommended to urgently establish a dedicated bioinformatic and data management centre specialized in marine microbial research. The centre should be designed as a specialised node of the EBI-ELIXIR infrastructure, to take full advantage of the existing capacities and competences in storage, processing and access to data.

The main tasks of the centre would be to:

- Process, analyse and integrate molecular data on the diversity and function of marine organisms with respect to the biotic and abiotic factors surrounding

them. Interoperability with oceanographic data and a high-level of standardization will guarantee a seamless data integration between researchers, national and European marine stations and EBI-ELIXIR;
- Build up specialist databases for marker genes and omics to support the analysis of biodiversity and function. All data should be curated by domain experts with respect to quality, consistency, interoperability and compliance to standards. The respective knowledge-enhanced databases will be a cornerstone for the sustainable analysis of marine molecular data;
- Develop controlled vocabularies, ontologies and data exchange formats in collaboration with experimental scientists, data providers and EBI-ELIXIR. This will enhance data annotation, electronic interoperability among databases and data exchange among users. Emphasis should be placed on data and quality management to allow sustainable data integration for marine ecosystems biology and biotechnology;
- In support of the above, offer training facilities and opportunities in bioinformatics and data management.

The proposed Marine Bioinformatics Centre will enable a new level of integrative and interdisciplinary marine research, allowing unprecedented insights into the complex interplay of marine organisms and their interaction with the environment. The centre should act as a mediator between marine researchers and the EBI-ELIXIR e-infrastructure to facilitate sustainable data management and knowledge generation in Europe.

Recommendation 4.
Promote interest in marine microbial research and improve training and education

Earth is a blue planet. With 70% of its surface covered by the ocean and 40% of the world's population living within 50 km of a coastline, our heritage, economy and well-being are inextricably linked with the marine environment. Marine microbes are the "gatekeepers" for the Earth System with an estimated contribution to global primary productivity of between 50% and 90%. It has become clear that marine microbes contribute fundamentally to all global cycles of energy and matter and it is evident that without their enormous power and potential, Earth would not be a living planet. Although these impressive facts are established, the proportion of research funding allocated to marine research and, in particular to marine microbial research, remains incredibly low. It is imperative that the distinct role of marine microbial communities in global change, sustainable supply of food and energy as well as human and environmental health is brought to the attention of all stakeholders from school children and students to the general public as well as industry, funders and policy makers.

The following key enabling actions should be supported as a matter of urgency:
- Support and promote open days, movies and lectures in public places and science educational facilities such as aquaria and museums. This should include integrated education material in schoolwork, and more targeted publications in specialized media and media campaigns towards policy makers;
- Develop specialized mobile exhibitions that evoke the imagination and interest of the public at large. These could include:
  – An exhibition ship that travels along the coast of Europe providing talks and lectures as well as hands-on experience to discover the microbial world in the marine environment;
  – An exhibition/education/training centre on marine microbial life made up from one or more containers with installed microscopes and display materials which can be transported by ship and by trucks from city to city.
- Create programmes at national, regional and pan-European level to improve graduate education. The future of life sciences in the 21st century is closely linked to the ability of scientists to develop and participate in interdisciplinary projects embracing skills and concepts from other disciplines. Hence, training the next generation of marine microbiologists must focus on the use of interdisciplinary and holistic approaches to solve a wide range of challenges specific to dealing with marine microorganisms and the marine environment. Desirable actions include:
  – Ensure that appropriate microbiology modules are included in all bio-science undergraduate educational programmes;
  – Initiate actions to ensure the participation of researchers from non-marine backgrounds in marine science and marine microbiology in particular, thus providing a growing pool of exceptional research talent for academic and industry needs;
  – Organize regular trainings or summer schools on marine microbiology subjects supported, for example, by the EU Framework Programme.

5.5 Implementing the recommendations

The four high-level recommendations of this position paper are designed in such a way that they are interdependent, complimentary and positively reinforcing. Creating better visibility and recognition for the importance of marine microbial diversity and associated research will assist in creating the necessary leverage for appropriate science policy measures and funding of relevant capacities and research activities. At the same time, the realization of a major pan-European research programme on marine microbiology will create a wealth of material and opportunities for outreach and education which will undoubtedly raise the profile of marine microbial research in Europe. It will also help to increase the appeal of this research field and attract the best students for a career in this fascinating domain.

The creation of a multidisciplinary European Centre for marine data management and analysis will be necessary to prevent the large scale implementation of omics technologies turning into a disruptive process (where the data production is faster than the speed users are able to make use of it). It will provide critical support and capacity to deal with the wealth of samples and genetic information that will come on-stream through the research in the framework of the pan-European research programme on marine microbiology. Likewise, material gathered during sampling campaigns would be preserved and made available to a larger community if the proposed European repository for cultivated collections and a reference library were to become a reality. The combined set of recommendations, when realized, will provide an enormous positive impact on European marine microbial research and provide the necessary push to improve Europe's competitiveness in this important research area.

A key consideration is the urgency with which action is required. By acting now to develop a coordinated, planned and properly resourced programme, as de-

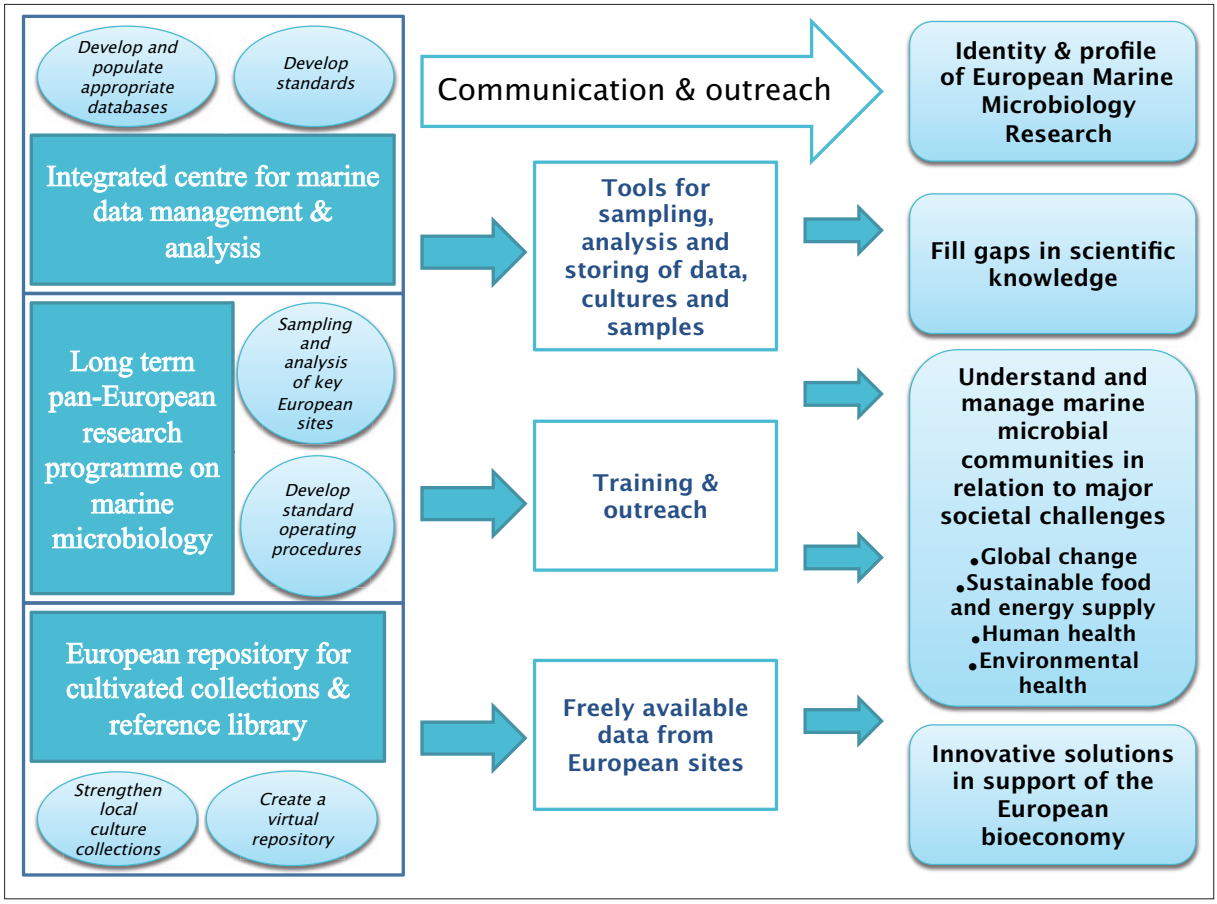


Figure 5.2. Schematic overview of main recommendations and their expected impacts in relation to addressing the key societal and scientific challenges

scribed above, we avert the damaging prospect of further disintegration and fragmentation of the European research effort which is inevitable under the "do nothing" scenario. Without this integrating framework and its key bioinformatics e-infrastructure, Europe will continue to slide in its position against other advanced research nations and, more importantly, opportunities to address key economic, social and environmental challenges will be lost. The European scientific community stands ready to work in harmony to advance European marine microbial research, knowledge and understanding. Let us all, as stakeholders in this knowledge, work together to deliver a true European Research Area in marine science and microbiology.

References

Arnaud-Haond S, Arrieta JM, Duarte CM (2011). Global genetic resources. Marine biodiversity and gene patents. Science 331:1521-1522.

Arrieta, JM, Arnaud-Haond S, Duarte CM (2010). What lies underneath: Conserving the oceans' genetic resources. Proc Natl Acad Sci USA 107:18318-18324.

Baldauf SL (2003). The deep roots of eukaryotes. Science 300:1703-1706.

Barton HN, Briggs DEG., eisen JA, Goldstein DB, Patel NH (2007). Evolution. New York, Cold Spring Harbor Laboratory Press, pp.120-123.

Blazejak A, Erséus C, Amann R, Dubilier N (2005). Coexistence of Bacterial Sulfide Oxidizers, Sulfate Reducers, and Spirochetes in a Gutless Worm (Oligochaeta) from the Peru Margin. Appl Environ Microbiol 71(3):1553–1561.

Brochier-Armanet C, Boussau, B, Gribaldo S, Forterre P (2008). Mesophilic crenarchaeota: Proposal for a third archaeal phylum, the Thaumarchaeota. Nature Reviews Microbiology 6(3):245–52.

DeLong EF, Karl DM (2005). Review Genomic perspectives in microbial oceanography. Nature 437:336-342.

Giovannoni S, Stingl U (2007). The importance of culturing bacterioplankton in the omics age. Nature Reviews Microbiology 5:820-826.

Glöckner FOG, Joint I (2010). Marine microbial genomics in Europe: current status and perspectives. Microbial Biotechnology 3:523-530.

Gómez-Consarnau L, Akram N, Lindell K, Pedersen A, Neutze R, Milton DL, González JM, Pinhassi J (2010). Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation. PLoS Biol 8(4): e1000358. doi:10.1371/journal.pbio.1000358.

King RD, Whelan KE, Jones FM, Reiser PGK, Bryant CH, Muggleton SH, Kell DB, Oliver SG (2004). Functional genomic hypothesis generation and experimentation by a robot scientist. Nature 427(6971):247-252.

Pedrós-Alió C (2006). Marine microbial diversity: can it be determined? Opinion. TRENDS in Microbiology 14(6). doi:10.1016/j.tim.2006.04.007

Sandaa RA (2009). Viral pool, environmental. In: Encyclopedia in Microbiology, third edition, Ed: Schaechter M. Section: Environmental Microbiology and Ecology. pp 553-567. Elsevier Ltd.

Sarmento Montoya JM, Vázquez-Domínguez E, Vaqué D, Gasol JM. (2010). Warming effects on marine microbial food web processes: how far can we go when it comes to predictions? Phil Trans Soc B 365:2137-2149.

Simon C, Daniel R (2011). Metagenomic analysis: Past and future trends. Applied and Environmental Microbiology 77:1153-1161.

Stratton MR, Campbell PJ, Futreal PA (2009). The cancer genome. Nature 458, 719-724.

Verberkmoes NC, Denef VJ, Hettich RL, Banfield JF (2009). Systems Biology: Functional analysis of natural microbial consortia using community proteomics Nature Reviews Microbiology 7:196-205.

Whitman WB, Coleman DC, Wiebe WJ (1998). Prokaryotes: The unseen majority. PNAS 95(12):6578-6583.

Zoetendal EG, Rajilić-Stojanović M, de Vos WM (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. Gut 57:1605-1615.

List of abbreviations and acronyms

AOB: Ammonia oxidizing bacteria
ARISA: Automated ribosomal intergenic spacer analysis
ASSEMBLE: Association of European Marine Biological Laboratories (EC FP7 research infrastructure initiative)
ATCC: American Type Culture Collection
ATP: Adenosine triphosphate
AWI: Alfred Wegener Institute for Polar and Marine Research
BGI: Beijing Genomics Institute
BODC: British Oceanographic Data Centre
CAMERA: Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis
CCN: Cloud condensation nuclei
CCY-NIOZ: Culture Collection Yerseke – Royal Netherlands Institute of Sea Research
cDNA: Copy- or Complementary DNA; refers to DNA synthesized from a messenger RNA (mRNA) template in a reaction catalyzed by the enzyme reverse transcriptase and the enzyme DNA polymerase
CLAMER: Climate Change and Marine Ecosystem Research (EC FP7 Project)
CoML: Census of Marine Life
CSA: Coordination and Support Action
CSIC: Spanish National Research Council
CTD: Conductivity-temperature-depth are standard oceanographic parameters measured by, among others, most research vessels, floats and automated underwater vehicles.
CWG-MB: Collaborative Working Group on Marine Biotechnology
DGGE: Denaturing Gradient Gel Electrophoresis
DMS: Dimethylsulfide
DMSP: Dimethyl sulfoniopropionate
DNA: Deoxyribonucleic acid; contains the genetic instructions used in the development and functioning of all known living organisms (with the exception of RNA viruses).
DNRA: Dissimilatory nitrite or nitrate reduction to ammonia
DOC: Dissolved organic carbon
DON: Dissolved organic nitrogen
DOP: Dissolved organic phosphorus
DSMZ: German Collection of Microorganisms and Cell Cultures
EBI: European Bioinformatics Institute (EBI)
EC: European Commission
EDMED: European Directory of marine environmental data
EDMERP: European directory of marine environmental projects
EiFEX: European Iron Fertilization Experiment
ELIXIR: European Life-Science Infrastructure for Biological Information
EMBL: European Molecular Biology Laboratory (EMBL)
EMBRC: European Marine Biological Resources Centre (EMBRC)
EMODNet: European Marine Observation and Data Network
EMP: Earth Microbiome Project
ENA: European Nucleotide Archive
EPS: Extracellular polymeric substances
ERA: European Research Area
ESF: European Science Foundation
ESP: Environmental Sampling Processor
ESTOC: European Station for Time series in the OCean
EU: European Union
EUROBIS: European Ocean Biogeographic Information System
EUROCORES: European Collaborative Research Scheme (developed and operated by the European Science Foundation)
EUROFLEETS: Towards an Alliance of European Research Fleets (EC FP7 Research infrastructures project)
EuroMarine: Integration of European Marine Research Networks of Excellence
FISH: Fluorescent In Situ Hybridisation
FP: Framework Programme
GCMD: Global change master directory
GEOSS: Group on Earth Observations
GMES: Global Monitoring for Environment and Security
GOS: Global Ocean Sampling
GPS: Global Positioning System
HABs: Harmful Algal Blooms
HERMES: Hotspot Ecosystems Research on the Margins of European Seas (EC FP6 Project)
HSV: Herpes Simplex Virus
ICES: International Council for the Exploration of the Sea
ICoMM: International Census of Marine Microbes
IMP: Integrated Maritime Policy
IMTA: Integrated Multi-Trophic Aquaculture
INSDC: International Nucleotide Sequence Database Collaboration
ISIS: In situ Iron Studies (consortium)
KEGG: Kyoto Encyclopedia of Genes and Genomes
MaCuMBA: Marine Microorganisms: Cultivation Methods for Improving their Biotechnological Applications (EC FP7 Project)
MAMBA: Marine Metagenomics for new Biotechnological Applications (EC FP7 Project)
Marbef: Marine Biodiversity and Ecosystem Functioning (EC FP6 Network of Excellence)
MAR-FISH: Microautoradiography-FISH
MBARI: Monterey Bay Aquarium Research Institute
MDs: Microbial Deliverables. The results of marine microbial research (data, information products, publications), and, in particular, marine microbial research products (samples, cultures and sequences) which are collectively called Microbial Deliverables (MDs) for the purpose of this position paper.
MEGAN: MEtaGenome Analyzer; a software package developed and used to analyze metagenomes
MEGX: is an ecologically oriented web portal that collects genome and metagenome data, displays it in a geo-referenced global system that combines the genetic data with ecological data obtained from the World Ocean Atlas and World Ocean Database.
MEOR: Microbial Enhance Oil Recovery
MGE: Marine Genomics Europe (EC FP6 Network of Excellence)
MG-RAST: Metagenomics RAST server; an automated analysis platform for metagenomes providing quantitative insights into microbial populations based on sequence data.

MicroB3: Biodiversity, Bioinformatics, Biotechnology (EC FP7 Project); aims to develop innovative bioinformatic approaches and a legal framework to make large-scale data on marine viral, bacterial, archaeal and protists genomes and metagenomes accessible for marine ecosystems biology and to define new targets for biotechnological applications.
mRNA: messenger RNA
MRSA: Methicillin-resistant *Staphylococcus aureus*
NASA: National Aeronautics and Space Administration
NCBI: National Center for Biotechnology Information
OIF: Ocean Iron Fertilisation
OMZ: Oxygen minimum zones
ORFs: Open reading frames
OTU: Operational taxonomic units
PANGEA Publishing network for geo-scientific and environmental data
PCR: Polymerase Chain Reaction
PESI: Pan-European species infrastructure
POC: Particulate organic carbon
POGO: Partnership for Observation of the Global Oceans
RADIALES: Oceanographic time series in northern Spain
RAS: Recirculating Aquaculture Systems
RAST: Rapid Annotation using Subsystem Technology
RCC: Roscoff Culture Collection
RFO: Research Funding Organisation
RNA: Ribonucleic acid
ROSCOP: Report of Observations/ Samples collected by Oceanographic Programmes
RPO: Research Performing Organisations
rRNA: ribosomal RNA
SAGE: Serial Analysis of Gene Expression
SeaDataNet: Pan-European Infrastructure for Ocean & Marine Data Management
SILVA: Comprehensive ribosomal RNA databases
SIP: Stable-isotope probing
SOFEX: Southern Ocean Iron Experiment
SOIREE: Southern Ocean Iron Enrichment Experiment
sRNA: small RNA
SST: Sea surface temperature
T-RFLP: Terminal Restriction Fragment Length Polymorphism
UNOLS: University-National Oceanographic Laboratory System
UV: Ultra Violet
VMDC: Flanders Marine Data Centre
WDC-MARE: World Data Centre for MARine Environmental sciences
WG MICROCEAN: Marine Board Working Group on Marine Microbial Diversity
WGA: Whole genome amplification
WORMS: World Register of Marine Species

Annex I. Participating authors

Contributing authors

Frank Oliver Glöckner *(Chair)*
Max Planck Institute for Marine Microbiology Bremen and Jacobs University Bremen, Germany
e-mail: fog@mpi-bremen.de

Adriano Bordalo
Institute of Biomedical Sciences, University of Porto, Portugal
e-mail: bordalo@icbas.up.pt

Josep M. Gasol
Institute of Marine Sciences, CSIC, Catalunya, Spain
e-mail: pepgasol@icm.csic.es

Francisco Hernandez
Flanders Marine Data Centre (VMDC), Belgium
e-mail: tjess@vliz.be

Matthias Labrenz
Leibniz Institute for Baltic Sea Research Warnemünde, Germany
e-mail: matthias.labrenz@io-warnemuende.de

Fergal O'Gara
BIOMERIT Research Centre, National University of Ireland Cork (UCC), Ireland
e-mail: f.ogara@ucc.ie

Paraskevi Pitta
Hellenic Centre for Marine Research, Institute of Oceanography, Greece
e-mail: vpitta@hcmr.gr

Ruth-Anne Sandaa
University of Bergen, Norway
e-mail: Ruth.Sandaa@bio.uib.no

Lucas J. Stal
Netherlands Institute of Ecology NIOO-KNAW, Netherlands
e-mail: L.Stal@nioo.knaw.nl

Elena Stoica
National Institute for Marine Research and Development "Grigore Antipa", Romania
e-mail: estoica@alpha.rmri.ro

Marta M. Varela Rozados
Spanish Institute of Oceanography (IEO) – Centro Oceanográfico da Coruña, SPAIN
e-mail: marta.varela@co.ieo.es

European Science Foundation

The European Science Foundation (ESF) is an independent, non-governmental organisation, the members of which are 72 national funding agencies, research performing agencies and academies from 30 countries.

Since its establishment in 1974, ESF, which has its headquarters in Strasbourg with offices in Brussels and Ostend, has assembled a host of organisations that span all disciplines of science, to create a common platform for cross-border cooperation in Europe.

ESF is dedicated to promoting collaboration in scientific research and in funding of research and science policy across Europe. Through its activities and instruments, ESF has made major contributions to science in a global context. ESF covers the following scientific domains:

- Humanities
- Life, Earth and Environmental Sciences
- Medical Sciences
- Physical and Engineering Sciences
- Social Sciences
- Marine Sciences
- Materials Science and Engineering
- Nuclear Physics
- Polar Sciences
- Radio Astronomy
- Space Sciences

www.esf.org

Printing: Drukkerij De Windroos NV (Beernem, Belgium)
May 2012
ISBN: 978-2-918428-71-8

Royal Netherlands Institute for Sea Research

Netherlands Organisation for Scientific Research

Marine Board
Wandelaarkaai 7
8400 Ostend • Belgium
Tel: +32 59 34 01 63
Fax: +32 59 34 01 65
info@marineboard.eu
www.marineboard.eu

May 2012 – Print run: 1000