MOLECULAR TOOLS FOR THE STUDY OF MARINE MICROBIAL DIVERSITY

L.K. Medlin, K. Valentin, K. Metfies and K. Töbe
Department of Biological Oceanography, Alfred Wegener Institute for Polar and Marine Research, Germany

R. Groben
Lake Ecosystem Group, Centre for Ecology & Hydrology, United Kingdom

Keywords: AFLP, algae, biodiversity, clone library, DGGE, DNA fingerprinting, DNA microarrays, FISH, Fluorescence In Situ Hybridization, microsatellites, molecular marker, molecular probes, oligonucleotide probes, PCR, phylogeny, phytoplankton, RAPD, RFLP, rRNA, solid-phase cytometry, SSCP, TGGE.

Contents

1. The Importance of Biodiversity Research in the Marine Environment
2. What Questions can be Answered Using Molecular Biology Techniques?
3. Evaluating Marine Biodiversity by Sequence Analysis and Fingerprinting Methods
   3.1. Sequence Analysis
      3.1.1. Which Genes to Select?
      3.1.2. How to Generate Sequence Data?
      3.1.3. Determining Biodiversity in an Environmental Sample by Sequence Analysis
      3.1.4. Analysing Sequences for Determining Phylogenies and Biodiversity
      3.1.5. DNA Barcoding
   3.2. Fingerprinting Methods
4. Analysis of Population Structure Using Molecular Markers
5. Molecular Probes for Identification and Characterisation of Marine Phytoplankton
   5.1. Introduction
   5.2. Probe Design
   5.3. Detection Methods
6. Conclusions
Acknowledgements
Glossary
Bibliography
Biographical Sketches

Summary

Marine photosynthetic microbial organisms are the major, sustaining components of ecosystem processes and are responsible for biogeochemical reactions that drive our climate changes. Despite this, many marine microorganisms are poorly described and little is known of broad spatial and temporal scale trends in their abundance and distribution. With new molecular and analytical techniques we can advance our knowledge of marine biodiversity at the species level to understand how marine biodiversity supports ecosystem structure, dynamics and resilience. We can then interpret environmental, ecological and evolutionary processes controlling and
structuring marine ecosystem biodiversity. With better analytical methods available, we can augment our understanding of biodiversity and ecosystem dynamics in especially the pico- and nano-fractions of the plankton as well as in the deep sea benthos, both of which are very difficult to study. Here we provide examples of new and long standing molecular tools for researchers in marine ecosystems to enable them to provide better, faster and more accurate estimates of marine biodiversity in the community using tools at the forefront of molecular research.

1. The Importance of Biodiversity Research in the Marine Environment

Understanding and preserving biodiversity was one of the most important global challenges for the past 20 years and will continue to be an important scientific issue during next decades. In 1999 the Association of Marine Science Institutes has recognised the need for a science plan for Europe to address the problems associated with a potential loss of biodiversity in the marine environment. This section on the importance of biodiversity research is a synopsis of the executive summary formulated for European research on marine biodiversity reflecting the joint opinions of scientists from the Association of Marine Science Institutes.

The global environment is experiencing rapid and accelerating changes, largely originating from human activity, whether they come from local requirements or from the more dispersed effects of global climate change. Widespread realisation that biodiversity is strongly modified by these changes has generated plans to conserve and protect biodiversity in many parts of the world that were heretofore subject to rampant savaging for natural resources (see also – Biodiversity: The Impact of Biotechnology). Adequate ecosystem functioning and therefore the continued use of the goods and services that ecosystems provide to humans depends on how biodiversity is perceived and preserved. Thus it follows that knowing and recognising biodiversity at all levels is an essential strategy for preserving biodiversity. Basic differences occur between terrestrial and marine ecosystems and the management of their individual biodiversity requires very different approaches. Generalisations concerning biodiversity patterns on both global and regional scales, the mechanisms that determine these patterns, and the consequences of biodiversity loss, are largely extrapolated from the terrestrial ecosystems, and many of these extrapolations are not applicable to the marine environment. Our understanding of marine biodiversity lags far behind that of terrestrial biodiversity, to such an extent that we do not have enough scientific information to design management plans, such as conservation and the sustainable use of coastal resources. Some of the fundamental differences between marine and terrestrial biodiversity include: the physical environment in the oceans is three dimensional, whereas on land it is only two-dimensional. The main marine primary producers are very small and usually mobile, whereas on land primary producers are large and stationary. Higher level carnivores often play key roles in structuring marine biodiversity and when exploited heavily as in overfishing, there are severe cascading downward effects on biodiversity and on ecosystem functions. This does not apply to terrestrial systems. Marine systems are more open than terrestrial and dispersal of species occurs over much larger ranges than on land. Life has originated in the sea and
thus is much older in the sea than on land. As a consequence, the diversity at higher
taxonomic levels is much higher in the sea and there are 14 indigenous marine animal
phyla, whereas only one phylum is unique to land. The sum total of genetic resources in
the sea is therefore inferred to be much more diverse in the sea than on land. Also on
average, genetic diversity within a species (i.e. below the species level) is higher in
marine than in terrestrial species.

Biodiversity must be evaluated at different scales: These are hierarchical levels (e.g.,
genetic, species, ecosystems), with spatial scales ranging from single samples to
regional and global, and temporal scales changing from short time intervals (days to
weeks) to long (years to decades). Threats to marine biodiversity and the consequences
of biodiversity loss or change operate at all of these scales. Results from research
conducted at any single level can lead to errors and unsupported expectations if
extrapolated to other levels. Biodiversity is more widely exploited in the sea than on
land: man commercialises over 400 species as food stocks from the marine
environment, whereas less species are utilised on land. Exploitation of marine
biodiversity is also far less regulated than that on land and amounts in the marine
environment to the hunting-gathering stage that humans abandoned on land about 10
000 years ago, but technology is becoming so advanced that many marine species are
now threatened and many are even extinct.

Marine organisms play pivotal roles in many biogeochemical processes that sustain the
biosphere, and provide a variety of goods and services that are essential to mankind’s
existence, including food production, assimilation of waste and regulation of the global
climate. Conservation efforts affect only marine reserves and specially protected areas
and the species they contain, which cover at best only a small part of the world’s marine
water. Thus, adequate functioning of marine systems depends in turn on biodiversity
and that fact dictates the need for a broader strategy in the management of biodiversity
than conservation alone can accomplish. Any biodiversity project must begin with
characterisation of the biodiversity as fully as possible (from genetic to ecosystem level)
in selected key (flagstone) habitats across broad geographical ranges. Compiling
comprehensive inventories at a few sites should do this and less comprehensive surveys
at a larger number of sites, using standardised methods and protocols. An important
example of a project that addresses these requirements is the Census of Marine Life, a
global network of researchers from over 70 countries that tries to answer the questions
“What lived in the oceans?”, “What lives in the oceans?” and “What will live in the
oceans?”. Molecular methods are an indispensable tool in answering those questions.

The world’s oceans cover 70 percent of the Earth’s surface and these areas are
dominated numerically by microscopic protists and prokaryotes. The marine
phytoplankton are, by definition, high dispersal taxa with large population sizes and are
major components of both these groups. The bulk of primary production in oceanic and
neritic waters involves these small photosynthetic organisms. Until recently most of our
knowledge about marine phytoplankton was derived from net samples and bulk process
measurements, such as chlorophyll a and 14C biomass estimates. However, previously
unrecognised groups (such as Prochlorococcus), size classes (the picoplankton < 3 µm)
and hidden biodiversity (new algal classes, e.g., Bolidophyceae and Pelagophyceae)
have been found by utilising whole water samplers and new analytical methods, e.g.,
flow cytometry, epifluorescence microscopy and HPLC (high pressure liquid chromatography). Surprisingly the picoplankton may contribute up to 90 percent of the photosynthetic carbon in certain areas. The picoeukaryotes and the cyanobacteria taxa *Prochlorococcus* and *Synechococcus*, whose importance in the open ocean oligotrophic ecosystems has only been discovered within the last 20 years are among this smallest size fraction of the marine phytoplankton. It has been shown that the eukaryote picoplankton are far more diverse than the prokaryote component.

We may question the accuracy of our knowledge about the genetic diversity of marine phytoplankton with these new revelations into phytoplankton biodiversity. In groups, especially the photosynthetic flagellates, where even α-level taxonomy is lacking, or in groups, such as the picoeukaryotes, where there are far too few morphological markers upon which to determine species identification, we soon realise that we probably know very little about their diversity. In addition we know virtually nothing about the population structure of the phytoplankton. It is likely to be very different from that on land because marine planktonic organisms live in an ever-changing three-dimensional environment. Many taxa may have little genetic structure over very large geographic areas. Further, recent evidence suggests that speciation and dispersal mechanisms in marine planktonic organisms may be very different from those on land. Thus, it is unlikely that generalisations about terrestrial plant diversity and population structure can be extrapolated to marine ecosystems.

2. What Questions can be Answered Using Molecular Biology Techniques?

The advent of molecular biological techniques has greatly enhanced our ability to analyse all populations, not just the phytoplankton. Their small size and paucity of morphological markers, the inability to bring many into culture, and the difficulty of obtaining samples for long term seasonal studies in open ocean environments has hampered our knowledge of phytoplankton diversity and population structure. Despite this, physiological/biochemical measurements have been used to infer the existence of significant genetic diversity within and between phytoplankton populations. These data have been used to speculate on hidden biodiversity and temporal and spatial structuring of genetic diversity or gene flow. Now molecular techniques can present a quantitative framework through which the diversity, structure and evolution of marine phytoplankton populations can be analysed, predictive models of the dynamics of ocean ecosystems formulated, and the idea of functional groups in the plankton proven.

Molecular analysis of phytoplankton population structure is behind other groups and has been usually inferred from physiological data determined from relatively few clones. This unfortunately is a very naive approach because many physiological measurements have shown that no single clone of any phytoplankton species can be considered truly representative of that species. The need to establish clonal cultures prior to genetic analysis and the inability to perform fine-scale sampling under most conditions are probably the overlying reasons why studies of phytoplankton population structure are perhaps 20 or more years behind those of other organisms. Isozyme analysis, performed for a few species, has revealed heterozygosity between some populations. In addition, fingerprinting analyses, such as RAPDs (random amplified polymorphic DNA) or multi locus probes, have shown that phytoplankton blooms are not clonal but are highly
diverse with isolates being related by geographic origin.

The interaction of a species with environmental parameters is influenced by the genetic diversity at the population level of a species. Spatial and temporal partitioning of genetic diversity will occur as these interactions structure the ecosystem. Such structuring has seldom been measured in the marine planktonic community and studies of genetic diversity are virtually non-existent in pelagic ecosystems. All evidence of geographically isolated populations would be erased if we continue to assume that marine organisms with high dispersal capacities are genetically homogeneous over their entire range. Support for this assumption has come mainly from phenotypic comparisons based initially on net phytoplankton biogeographic studies and later on isozyme studies. Reason why studies of phytoplankton diversity and population structure have lagged behind those of other organisms is because of their small size and the lack of morphological markers, and the ability to bring into culture only a small part of the known biodiversity. The lack of knowledge of their breeding systems makes genetic or demographic studies difficult. Logistical problems of collecting samples for long term seasonal studies in open ocean environments or for doing fine-scale sampling are additional reasons.

Another issue is whether adequate sampling strategies can be employed for phytoplankton populations to address spatial and/or temporal genetic variation questions. Pre-established cruise tracks may make the sampling of oceanic populations only possible at depth rather than in a hierarchical grid-like fashion that may be needed for population studies. A lack of knowledge about current regimes in the study area may also bias sampling strategies if samples are unknowingly taken from two water masses. At present most genetic studies must rely on clonal cultures for their analyses. These single-cell isolations are made from natural populations and can be difficult to perform at sea. The selective survival of only 10 – 30 percent of clones from natural populations may mean that the range of genetic diversity determined from a bank of clonal isolates may not be a true reflection of the genetic diversity in the original population and may not be adequate for the level of genetic diversity being addressed. In many algal groups, life histories are incomplete, and if the algae undergo sexual reproduction during culturing, then this may also affect the type of genetic analysis performed (see also – Algal Cell Culture)

As far back as the mid-seventies Doyle hypothesised that planktonic algae must consist of a multitude of competing genotypes, but this study was largely ignored and it was assumed that planktonic taxa may have little genetic structure over very large geographic areas. Marine planktonic organisms are not sessile but are constantly mixed by currents and waves and it was assumed that highly dispersed organisms at the mercy of these factors have no trace of genetic structure. With the possibilities of nucleic acid methods, however, these views on the absence of genetic structure in the marine phytoplankton have been seriously challenged. Genetic structure and physical, spatial partitioning within biogeographic regions are now known. The idea of a single globally distributed species is no longer believed, nor is the idea of temporal stasis. Temporal genetic changes can often be greater than spatial changes or changes between species. This may very well apply to bloom populations. The rate of genetic change can and does occur on ecological time scales. Reasons for this are unclear but such changes
may play a role in determining how local adaptations and speciation can occur in apparently homogeneous populations. The concept of a 'super species' with the ability to exploit a wide spectrum of environmental conditions may lay the groundwork for temporal genetic change.

Much of our limited knowledge about phytoplankton genetic diversity stems from the difficulty of finding polymorphic markers for ecological genetic studies. Isozymes, the molecular genetic markers used in early studies, evolve so slowly that closely related populations appear identical. This fact has undoubtedly propagated the early ideas of the absence of genetic diversity in marine phytoplankton. The use of high resolution molecular marker techniques *sensu lato* circumvents these problems and has thus opened areas previously considered intractable.

Plastid and flagellar apparatus characteristics are the features that define most phytoplankton classes, making them monophyletic taxa, but some surprises have been revealed by molecular analyses. For example, the Euglenophyceae, once thought to be related to the Chlorophyceae, are shown to be a very early eukaryotic radiation and not part of the major eukaryotic radiation called the crown group radiation. The kingdom Chromista did contain the bulk of marine eukaryotic phytoplankton taxa, e.g., the Heterokonta, Haptophyta, and Cryptophyta. But this kingdom is now recognised as a polyphyletic taxon. Molecular analyses based on total evidence (both morphological and molecular data from the rRNA data set) continue to reinforce the clear separation of the haptophyte from the heterokont algae whereas those based on many other genes have distanced the cryptophytes from both the heterokonts and the haptophytes. A fourth group, the Chlorarachniophyceae, were also formerly placed in the Chromista but are now shown to be clearly related to the foliose amoeba. Clearly the Kingdom Chromista is an idea whose time has past. New algal classes have been recognised from molecular analyses, e.g., Pelagophyceae and Bolidophyceae. Molecular analyses have now reached a general consensus that there are eight major groups of eukaryotes, one of which is the Chromalveolates, which contains the Chromista and the Apicomplexa. Concatenated plastid genes link all of the chromists to a single evolutionary event but host genes still separate the haptophytes and the cryptophytes from the other chromist and from the aforesaid eight major groups.

Molecular techniques have changed systematics dramatically at the genus and species level, showing polyphyletic and paraphyletic lineages across many algal groups, not just the phytoplankton. The most dramatic upheavals have come in groups with few morphological markers, and where morphological species definitions have been too broad. The prochlorophytes, *Chlorella*, *Chlamydomonas* and *Chrysochromulina* are all recognised as polyphyletic taxa. But even in groups with good morphological markers, e.g., *Skeletonema* and *Cryptocodonidium*, (cryptic) sibling species have been found; others not so easy to differentiate at the species level, e.g., *Phaeocystis*, also contain cryptic species. Nearly all cosmopolitan species in the micro size class have been shown by molecular techniques to be composed of multiple species. As a consequence, taxonomic revisions are under way in many phytoplankton groups, but likely cryptic species with no morphological distinction and only genetic distinction have to be accepted in all groups.
Molecular tools in general offer the possibility to estimate biodiversity at all levels, e.g., kingdom/class/family/species level, in a comparatively small environmental sample. In some cases even a few millilitres of seawater may be enough. Moreover, some of the techniques are very sensitive, e.g., offer the possibility to detect single cells in a sample. Depending on the question(s) being asked the molecular tools to answer them differ greatly. One may wish to detect as many species as possible in a given sample. In this case the establishment of an rDNA clone library with subsequent sequencing of as many clones as possible can uncover the biodiversity in the sample in great detail. General assessment of comparative biodiversity in a larger number of samples can be achieved with DNA fingerprinting methods based on denaturing or temperature gradient gel electrophoresis (DGGE, TGGE) or single strand conformation polymorphisms (SSCP). Presence or absence of a known species can be monitored with species-specific probes using chemiluminescent detection and DNA dot blot techniques or, more sophisticated, with fluorescent in-situ hybridisation (FISH). Distinction of individuals at the family or even species level can be obtained using highly variable molecular markers such as ITS sequences (inter-transcribed spacer) or microsatellites. Finally per se non-molecular techniques like flow cytometry that have already been used in the "premolecular age" can be combined with DNA techniques (staining of the nucleus, hybridisation with fluorescence-labelled specific oligonucleotide probes) to distinguish and quantify species in environmental samples.

In general, molecular techniques have some significant advantages over traditional methods:

- Only very small samples (in the range of millilitres up to a litre) are required for most analyses.
- Sensitivity of many methods is very high, e.g., enabling the researcher to detect even single specific cells among thousands of others.
- Dead or non-culturable cells can be analysed.
- Species-specific data (such as sequences) can be obtained without the need to culture or even isolate a species.

As with all methods, molecular ones also contain certain biases. The harvesting of cells through filtration or centrifugation may be harmful for fragile organisms, which thus may escape the analysis. For many techniques the lysis of organisms with subsequent isolation of DNA is a prerequisite. Both steps may not be equally effective in all organisms. In PCR-based approaches biases are evident concerning the choice of (universal) primers, PCR conditions (e.g., the amount of DNA or primers used, annealing temperature, cycle number etc.), machines or enzymes used etc. The copy number of genes of interest (mostly ribosomal RNA genes) differ greatly among various organisms. If cloning steps are involved, then the choice of vectors, enzymes or bacterial strains may be relevant. Hybridisation experiments are susceptible to hybridisation conditions (temperature, salt concentration, time) or base composition and subsequent detection of fluorescence may be hampered by autofluorescence. All the former is especially important when absolute quantification of results is desired. In general we advise the same caution when interpreting the results of molecular methods as for all other methods. Results are not more reliable because they come from a “molecular” approach rather than a ”classical” one.
In the following we will summarise and briefly explain a variety of techniques currently being used or under development. We also try to estimate the advantages and shortcomings of such methods.

TO ACCESS ALL THE 33 PAGES OF THIS CHAPTER, Visit: http://www.eolss.net/Eolss-sampleAllChapter.aspx

Bibliography


Census of Marine Life Http://www.coml.org/
[Global network of researchers with the aim to assess and explain the diversity, distribution, and abundance of marine life in the oceans.]

Consortium for the Barcode of Life. Http://barcoding.si.edu/index_detail.htm
[International initiative devoted to developing DNA barcoding as a tool for scientific research on the taxonomy of plant and animal species.]


Department of Genetics, University of Washington, USA.
Http://evolution.genetics.washington.edu/phylip/software.html
[A list of nearly 200 different software packages for phylogenetic analyses with comments and links for downloads. All programs mentioned in this chapter are listed here.]

[Good overview over different types of molecular markers and their use. Concentrates on forest genetics, but a lot of information is also of interest for a broader scientific community.]


National Center for Biotechnology Information, National Institute of Health, USA.


[This server offers free use of sequence comparisons by BLAST and access to databases like GenBank.]


[Overview of the range of biodiversity that has been reported in wide variety of marine habitats from the plankton to the deep-sea benthos]


Ribosomal Database Project II, Center for Microbial Ecology, Michigan State University, USA.

Http://www.cme.msu.edu/rdp/html/index.html

[This website provides ribosome related data services, including online data analysis, rRNA derived phylogenetic trees, and aligned and annotated rRNA sequences.]


Biographical Sketches

Linda Medlin received her Ph.D. at Texas A&M University in 1983 in marine botany. Her formal training lies primarily in the taxonomy and ecology of marine phytoplankton, with strong emphasis on the diatoms. Within the last ten years she had incorporated molecular techniques into a broad study of the phylogeny of some of the major algal classes, such as the diatoms, the dinoflagellates and the haptophytes. More recently she had begun in-depth study of the diatom phylogeny using rRNA sequencing and an assessment of their use in calibrating a molecular clock using the 18S rRNA molecule. Further development in the area of molecular clocks in her laboratory has involved the calculation of the emergence of several plastid genes to estimate the timing of secondary endosymbiosis events. Dr. Medlin's current research program also addresses the issue of understanding genetic diversity at the population and species levels in the marine phytoplankton using molecular techniques. They are making hierarchical probes with end goal of making DNA-microchips for their use in estimating biodiversity from marine samples and for making early warning systems for toxic algae.

Klaus Valentin studied chemistry and biology at the Justus Liebig University, Germany and finished his Ph.D in 1990. He worked as a Postdoc at the Dept. of Botany of the University of Washington in Seattle, USA and at the Justus Liebig University. Since 2000 he is at the Alfred-Wegener-Institut in Bremerhaven, Germany. His research interests were focused on the molecular biology and evolution of the plastid, his present work concerns the molecular analysis of picoplankton biodiversity in the oceans.

Katja Metfies studied biology at the University of Osnabrück and Hanover. She received her PhD in genetics at the University of Cologne in 2000. The scientific focus of her PhD was the investigation of transcriptional regulation in yeast. Since 2001 Katja Metfies works as a Postdoc at the Alfred Wegener Institute for Polar and Marine Research developing and using molecular probe based methods for the assessment of phytoplankton biodiversity.

Kerstin Töbe studied biology at the University of Bremen, Germany and finished her PhD in 2003 at the Alfred Wegener Institute for Polar and Marine Research in Bremerhaven, Germany. Her PhD thesis dealt with the investigation of bacteria and dinoflagellates interactions in paralytic shellfish poisoning. She is Postdoc at the Alfred Wegener Institute and works with molecular probes to detect toxic and non-toxic microalgae with Fluorescence in situ hybridisation in combination with solid phase cytometry.

René Groben studied biology at the University of Hannover, Germany, and finished his Ph.D. in 1997.
His Ph.D. thesis dealt with the development and application of microsatellite markers in sugar beet and in this context he also compared this method to other marker techniques. He worked as a Postdoc at the Alfred Wegener Institute for Polar and Marine Research developing molecular probes for a broad range of phytoplankton taxa with an emphasis on toxic species and investigates phytoplankton biodiversity using these probes and different kinds of molecular markers. Currently, he is a Senior Research Officer at the Centre for Ecology & Hydrology in Lancaster, U.K., were he applies molecular methods in studying phytoplankton diversity in freshwater lakes in correlation to environmental conditions.