Marine microbial ecology in a molecular world: what does the future hold?

DAVID A. CARON

Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway, AHF 301 Los Angeles, CA 90089-0371, USA. E-mail: dcaron@usc.edu

SUMMARY: Advances in genetic and immunological approaches during the last few decades have transformed medicine and biomedical research. The human genome and the genomes of numerous model organisms are now fully sequenced. Initial exploitation of this wealth of genetic information has begun to revolutionize research on these species, and the applications derived from it. Progress in understanding the ecology of microorganisms (including marine taxa) has followed closely on the heels of these advances, owing to the tremendous benefit afforded by major technological advances in biomedicine. Through the application of these novel approaches and new technologies, marine microbial ecology has moved from a minor footnote within marine biology and biological oceanography during the 1950s and ‘60s to the focus of much of our present interest in the ocean. During the intervening half-century we have learned a great deal regarding the overall abundances, distributions and activities of microorganisms in the sea. Recognition of the extraordinary diversity of marine microbes, the predominant role that they play in global biogeochemical processes, and the potential for natural or engineered microbial products to benefit humankind, has placed marine microbes in the spotlight of both scientific and popular attention. Our fascination with these minute denizens of the ocean is not likely to wane anytime soon. Recent studies have indicated that we still know relatively little about the breadth of microbial diversity in marine ecosystems. In addition, many (most?) of the predominant marine microbial forms in nature have not yet been brought into laboratory culture. Thus, our knowledge is still rudimentary with respect to the spectra of biochemical, physiological and behavioral abilities of these species, and the study of marine microbes will remain a major focus of investigations in marine science well into the foreseeable future. As a large cadre of researchers moves headlong into this work, we can expect many new discoveries and more paradigm shifts regarding the composition and function of marine microbial communities.

Keywords: microbial ecology, bacteria, protists, ecogenomics, molecular taxonomy.

RESUMEN: ECOLOGÍA MICROBIANA MARINA EN UN MUNDO MOLECULAR: ¿QUE NOS AUGURA EL FUTURO? – Los avances en los campos de la genética y la inmunología durante las últimas décadas han transformado las investigaciones de medicina y biomedicina. El genoma humano y el de muchos otros organismos modelo han sido completamente secuenciados. La explotación de esta importante información genética ha empezado a revolucionar la investigación de estas especies y las aplicaciones derivadas de ésta. El progreso en el conocimiento de la ecología de los microorganismos (incluyendo los marinos) ha seguido muy de cerca los avances tecnológicos en biomedicina. La aplicación de nuevas aproximaciones y nuevas tecnologías ha permitido que la ecología microbiana haya pasado de ser una nota a pie de página a ser el tema principal en muchos de los estudios en el océano. Durante los últimos 50 años hemos aprendido aspectos acerca de las abundancias, distribuciones y actividades de los microorganismos marinos. El reconocimiento de la extraordinaria diversidad de los microbios marinos, la importancia de su papel en los ciclos biogeoquímicos y las posibles aplicaciones de los productos fabricados por los microorganismos en beneficio de la humanidad, ha colocado los microbios marinos en el punto de mira tanto de los científicos como del público en general. Nuestra fascinación por estos seres diminutos es posible que prevalezca mucho tiempo. Estudios recientes han indicado que todavía conocemos relativamente poco acerca de la gran diversidad de microorganismos en sistemas marinos. Además, muchos (la mayoría?) de las formas predominantes en la naturaleza todavía no se han logrado cultivar en el laboratorio, lo que nos indica que nuestro conocimiento es todavía rudimentario en relación con las habilidades bioquímicas, fisiológicas y de comportamiento de estos microorganismos. Así pues, de forma previsible, el estudio de los microbios ocupará una gran parte de las investigaciones en ciencias marinas en el futuro. A medida que aumenten el número de investigadores dedicados a estos estudios, se vaticina muchos descubrimientos nuevos, así como nuevos paradigmas relacionados con la composición y función de las comunidades microbianas marinas.

Palabras clave: ecología microbiana, bacteria, protistas, ecogenómica, taxonomía molecular.
INTRODUCTION

The evolution of a modern paradigm

The study of marine microorganisms is undergoing a period of explosive growth and rapidly changing paradigm. This situation is due to the central role that these species play in energy flow and elemental cycling, a conclusion that has developed gradually during the past quarter century from improvements in our knowledge of the abundances, distributions and activities of microbial assemblages in the ocean. Marine microbiological studies up to the mid-1970s were fundamental in establishing the importance of high abundances of bacteria in seawater (Zobell and Feltman, 1938; Jannasch and Jones, 1959; Jannasch, 1966; Hobbie et al., 1972; Francisco et al., 1973), the preponderance of ‘minute’ phytoplankton among the primary producers (Malone, 1971), and the important roles played by heterotrophic unicellular eukaryotes (heterotrophic protists, aka protozoa) as links in microbial food webs and as agents of decomposition and nutrient remineralization (Johannes, 1965; Fenchel, 1969; Fenchel, 1970). Nevertheless, mathematical models of marine ecosystems up to that time typically did not include microorganisms (other than large phytoplankton) as major facets of marine food chains (Steele, 1974). This landscape changed dramatically in subsequent years, beginning with what many researchers view as the seminal paper in marine microbial ecology (Pomeroy, 1974), a review which formalized our knowledge up to that time regarding the ecological roles of microorganisms in oceanic ecosystems.

Improvements in microscopy during the 1970s to early 1990s, the application of new visualization technologies such as flow cytometry, the use of radioactive compounds as tracers to examine the pathways and rates of flow of elements and energy within marine food webs, and extensive physiological studies of cultured species of bacteria and protists (algae and protozoa) provided a wealth of information for expanding the newly emerging paradigm outlined by Pomeroy. These later studies established some of the basic patterns and constraints on the growth, consumption and elemental cycling of microorganisms in the ‘microbial loop’. They lead to a number of formalizations of their various roles over the next few decades, and thus a clearer understanding of the activities of ‘ecological groups’ or ‘guilds’ of microbial taxa (Sieburth et al., 1978; Azam et al., 1983; Fenchel, 1988; Sherr et al., 1988; Sherr and Sherr, 1988; Fasham et al., 1990; Ducklow, 1991; Caron and Finlay, 1994; Ducklow, 1994; Sherr and Sherr, 1994). This box-model approach has been useful for placing microbial populations into context with other marine taxa. These various conceptualizations have continued to undergo refinement as new discoveries and new insights have taken place.

New discoveries and breakthroughs in microbial ecology have typically arisen through the development and application of new technologies and approaches. This correlation is due to the strongly technique-limited nature of this field. For example, the application of epifluorescence microscopy to the examination of natural water samples was instrumental in establishing the high and relatively constant abundances of bacteria that exist in seawater (Jannasch and Jones, 1959; Francisco et al., 1973; Hobbie et al., 1977), and later applications of image analysis provided more information regarding size and biomass (Sieracki et al., 1985; Bjørnsen, 1986). Epifluorescence microscopy also played a pivotal role in the discovery of the presence of ubiquitous populations of chroococcoid cyanobacteria such as *Synechococcus* spp. in the world ocean (Johnson and Sieburth, 1979; Waterbury et al., 1979). The extension of epifluorescence microscopy to the study of protists in the 2-20 μm size class documented the important contributions of minute eukaryotic algae and heterotrophic flagellates in the plankton, and viruses in the <0.2 μm size class (Proctor and Fuhrman, 1991; Fuhrman and Suttle, 1993; Sherr et al., 1993). The application of flow cytometry to the analysis of oceanic waters provided evidence for the ubiquity of the chroococcoid cyanobacterium *Prochlorococcus* (Chisholm et al., 1988; Olson et al., 1990) and picoeukaryotic algae (Li et al., 1983; Li et al., 1992). The use of natural or surrogate fluorescently-labeled prey (Sherr et al., 1987; Rublee and Gallegos, 1989) to examine the feeding activities of phagotrophic protists was instrumental in establishing the major trophic relationships among microbial consumers (Sherr and Sherr, 1994; Caron, 2000a), as well as identifying mixed phototrophic/phagotrophic nutrition among many protistan taxa (Sanders, 1991; Stoecker, 1998; Caron, 2000b). Each of these findings has improved our understanding of the diversity, living biomass and dynamics of specific kinds of microbes in natural waters and forced significant shifts in the existing paradigm of carbon and energy flow within pelagic communities.
Recent progress on prokaryotes

Starting in the mid-1980s and into the early 1990s, approaches and methodologies within the field of molecular biology began to make their way into investigations of the diversity of natural assemblages of planktonic microorganisms (Pace et al., 1985; Olsen et al., 1986; Amann et al., 1990b; Giovannoni et al., 1990a; Giovannoni et al., 1990b; Fuhrman et al., 1993). These early studies were focused primarily on free-living bacteria. Prior to this period, the composition of these bacterial communities was known only through strains isolated and cultured from natural seawater samples. For the first time, genetic approaches enabled an analysis of natural microbial communities without the potential selectivity that laboratory enrichment and culture might impose. The application of single-cell probing technologies that paralleled and followed these initial characterizations of bacterial community composition allowed the first phylogenetic/taxonomic classification of individual bacterial cells in natural seawater samples (Giovannoni et al., 1988; DeLong et al., 1989; Amann et al., 1990a; Amann et al., 1990b; DeLong, 1991; Amann et al., 1995).

These early studies immediately changed the existing conventional wisdom regarding the breadth of prokaryote diversity and the identity of the most ecologically important taxa in the ocean. Unexpectedly, natural bacterial assemblages were shown to be composed of phylotypes (16S rDNA sequences) of taxa that were not present in extant culture collections of marine bacteria. This finding challenged the view that culture collections of the time captured the true diversity of natural bacterial assemblages, and it brought into question the validity of physiological information gleaned from existing culture collections for predicting bacterial processes taking place in the ocean. Not unexpectedly, these findings have spawned a new generation of culture approaches that hold promise for culturing the ‘unculturable’ bacteria of the ocean (Connon and Giovannoni, 2002; Rappe and Giovannoni, 2003; Stevenson et al., 2004).

Genetic analyses of prokaryotes have yielded other surprising findings. Phylogenetic analyses of gene sequences in the early 1990s indicated the widespread occurrence of microbial taxa that were only distantly related to the bacteria and supported the ‘Three Domain’ framework for organizing all living organisms (Woese et al., 1990). Subsequent characterizations demonstrated unexpectedly high abundances of Archaea throughout the world ocean (DeLong, 1992; Fuhrman et al., 1992; Karner et al., 2001). More recently, the ubiquity of highly unique and unexpected compounds such as proteorhodopsin (Beja et al., 2001), and physiological diversity such as the potentially widespread occurrence of aerobic anoxygenic phototrophy among prokaryote taxa (Kolber et al., 2000; Beja et al., 2002; Oz et al., 2005) have stimulated reevaluations of the fundamental trophic modes of many microbes in the sea (Karl, 2002).

Recent progress on eukaryotes

DNA sequence information has made equally significant contributions to the biology of eukaryotic (protistan) microbes. Historically, protistan species have been described based on morphological features of the cell (general size and shape, flagellation/ciliature, pigmentation, mineralized structures, etc.). Morphology remains the ‘gold standard’ for species identification at present, but it does have shortcomings. For example, minute protists often possess few distinctive morphological characters for species identification (e.g. many species <10 µm in size), while some taxa have variable or amorphous cell size or shape (e.g. amoebae). Many taxa possess multiple life stages that manifest different morphologies (e.g. cysts vs. vegetative cells). Due to these limitations and complications, genetic analyses have proven useful as supplemental taxonomic tools and for helping to establish the evolutionary relationships among protistan taxa (Sogin et al., 1986; Lipscomb, 1989; Sogin, 1989; Sogin, 1991; Schlegel, 1994), especially where sufficient morphological characters have been lacking (Lim et al., 2001; Sims et al., 2002) or in the face of conflicts arising from misleading or ambiguous morphological characters (Edgcomb et al., 2002).

Studies of natural protistan assemblages using molecular biological approaches emerged later than the first attempts to characterize bacterial/archaeal assemblages, but the former studies have increased rapidly and have resulted in equally significant changes in our perception of protistan community structure. In addition to the difficulties of species identification noted above, complicated procedures for collection, fixation and/or staining used to visualize diagnostic taxonomic characters have presented major hurdles for ecologists attempting to describe the composition of natural protistan assemblages. Genetic analyses provide one possible way...
Diversity surveys of microbial eukaryotes using small subunit ribosomal RNA genes (rDNA 18S sequences) have recently reported tremendous previously-undocumented diversity (Díez et al., 2001; López-Garcia et al., 2001; Moon-van der Staay et al., 2001; Edgcomb et al., 2002a; Habura et al., 2004). Many of the sequences retrieved from environmental samples have been novel phylotypes of taxa that are not present in culture collections. Some phylotypes may represent new diversity at the phylum or kingdom level (Dawson and Pace, 2002; Massana et al., 2002; Stoeck et al., 2003). Moreover, microbial eukaryote community structure is complex, with numerous rare phylotypes that are highly responsive to environmental conditions (Countway et al., 2005). These findings have stimulated efforts to further characterize these taxa and bring them into laboratory culture, an effort that has been analogous to new approaches to culture the ‘unculturable’ bacteria (Guillard, 1973; Guillou et al., 1999a; Guillou et al., 1999b; Massana et al., 2002).

In addition to diversity analyses, a major focus for the application of molecular biological approaches to protistan ecology has been the acquisition of accurate and quantitative information for ‘species of interest’. While molecular analyses of natural prokaryote assemblages have focused largely at the community level, autecological studies of individual protistan taxa have received an equivalent amount of attention as studies of protistan community diversity. Genetic methods for the identification of species of ecological, economic or human health interest have begun to emerge. These include methods for the study of harmful bloom-forming algae (Scholin et al., 1996; Bowers et al., 2000; Coyne et al., 2001; Popels et al., 2003), potential human pathogens (Gast and Byers, 1995), uncultivated picocyanobacteria (Massana et al., 2002) and species that may play important ecological roles in nature (Lim et al., 1999).

Paradigm interrupted

Collectively, recent findings from these initial attempts during the late 20th and early 21st century at applying molecular biological methods to characterize marine microbial communities have brought the field full circle, back to a ‘discovery phase’ that oceanography experienced more than a century ago. Early oceanographic exploration of the late 19th century was largely descriptive as each new trawl, plankton net tow or grab sample yielded new species and new insights into marine community structure. Although well beyond its ‘infancy’, microbial ecology repeats that pattern today, using the tools of modern molecular biology to plumb the depths of microbial diversity, to establish a functional understanding of the meaning of this diversity, and to place that information within the larger context of ocean biology and biogeochemistry. In doing so, microbial ecology has clearly stepped from the shadows to the forefront of marine research.

A BRIGHT FUTURE FOR MARINE MICROBIAL ECOLOGY

Predicting the future of a highly active scientific field is an entertaining but risky undertaking. Progress is typically non-linear and research directions often change rapidly. That being said, some of the general trajectories (and probable break-throughs) within the broad spectrum of marine microbial ecology are apparent at this time. Just as sequencing of the human genome has set the stage for some obvious directions in biomedical research, the application of genetic approaches to the study of marine microbes has set the stage for breakthroughs in our understanding of how these organisms function and how they interact with other microbes in the ocean (Fig. 1).

The ‘-omics’ revolution

DNA sequencing capabilities of high-throughput facilities have advanced at a dizzying pace in recent years. Sequence databases are expanding rapidly while the ability to obtain sequence information for any given organism is becoming almost trivial. The time required to completely sequence the genome of an individual organism has been reduced to days. In addition, metagenomic (environmental) data is beginning to amass the summed genomic information present in a wide variety of marine ecosystems. The ‘downstream’ expression of that genomic potential (i.e. the proteome and metabolome) will further add to the data available on microbial abilities and activities, but also greatly complicate the process of interpretation of the genetic potential of microbial genomes. This immense amount of infor-
Information is as overwhelming as it is enticing. We have reached the point where our ability to collect data exceeds our present ability to interpret it properly. Therefore, one means of discovery in the immediate future will be attained through the derivation of effective ways of analyzing the enormous amount of information becoming available, and devising approaches to exploit that information to gain knowledge of how microbial communities are structured and how they function.

**Taxon-directed genomics**

Sequencing of cultured species of microbes has proceeded exponentially during the last decade. A few hundred genomes are now complete or in progress at this time. Many of the most recently sequenced species are prokaryotic and eukaryotic marine microbial taxa (Derelle et al., 2002; Palenik et al., 2003; Rocap et al., 2003; Armbrust et al., 2004; Moran et al., 2004). At the present rate, many hundreds of genomes of cultured microorganisms may be sequenced within the next decade, and that information made available to the scientific community. Studies of these species are in such an early stage that we can only speculate on the likelihood of paradigm shifts as a result of information that will be mined from these genomes. Nevertheless, a few ‘low-lying fruit’ have already been plucked from these databases. For example, these studies have indicated unexpected metabolic capabilities among bacteria and eukaryotic phytoplankton (Armbrust et al., 2004; Moran et al., 2004), unique adaptations to differing light and nutrient regimes by oceanic cyanobacteria (Palenik et al., 2003; Rocap et al., 2003), and the possible evolutionary and ecological consequences of genome size in the cyanobacterium *Prochlorococcus* (Rocap et al., 2003). The future of this area of study seems boundless. Little (actually nothing) is known of the function of more than a quarter of the genes sequenced. This situation must and certainly will change in the future.

Genomic studies of individual taxa hold the promise of truly understanding the influence of the sur-
rounding environment on a microbe’s activities. It is important to recognize that the genetic potential contained in a genome does not necessarily equate to the realized activity or function of the microbe in nature, but it is a starting place. Combined with approaches such as macro/microarrays that will use genomic information to ‘interrogate’ cells or communities under natural or imposed conditions, these types of approaches will bring an experimental aspect into ecological/physiological studies that will become an important tool for understanding the interactions between a microorganism and its chemical, physical and biological environment. Manipulative studies in which an organism is exposed to different environmental conditions will elicit an understanding of the biochemical pathways or physiological responses used by specific microorganisms. Methodologies emerging at this time will soon allow these interrogations to take place on living single cells (Lidstrom and Meldrum, 2003). Such ‘life-on-a-chip’ approaches will provide unprecedented information on the response of cells to stimuli. ‘Whole community’ manipulations will allow queries about the status and response of specific processes in a natural community (Jenkins et al., 2004). Genomic studies also pave the way for designing bioengineered microbes to serve as biosensors (Simpson et al., 2001), or to direct biochemical processes such as the remediation of polluted environments (Valls and de Lorenzo, 2002).

Although feasible and ultimately desirable, complete genome sequencing is still beyond the capabilities of most individual research laboratories. Government and private support specifically for this task will continue to be essential, and a waiting period will continue to be required for many genome projects. Alternatives to complete genome sequencing are being applied to address these issues and to provide new and highly sensitive methods to query molecular processes operating within cells. The use of expressed sequence tags (ESTs), massively parallel signature sequencing (MPSS) and other emerging approaches offer the possibility of interrogating microbes for gene function (Brenner et al., 2000; Scala et al., 2002). These approaches are particularly appropriate for microbial species with very large genomes that would still require major sequencing efforts (e.g. many dinoflagellates).

**Community-directed genomics**

Genomic work at the community level undoubtedly will take the most time to fully resolve, but will also provide gains that cannot be fully predicted at this time. Some level of ‘undirected’ discovery is appropriate and necessary at this time because we know so little of the overall composition of natural microbial communities. Metagenomic programs (also referred to by a variety of epithets such as ‘ecogenomics’ or ‘environmental genomics’) that are presently underway epitomize this approach with their overarching goal of assaying the total microbial genetic diversity within an ecosystem (Stein et al., 1996; Beja et al., 2000; Rondon et al., 2000; Cary and Chisholm, 2001; Clark et al., 2004; Riesenfeld et al., 2004; Tyson et al., 2004). Most microbial metagenomic studies appear to lack a clear focus or direction. That is meant to be an observational statement rather than a criticism and is due, in part, to the fact that we do not yet have a good appraisal of the extent of the microbial diversity present in most ecosystems (Thompson et al., 2005) or the significance of that diversity for overall community function. Many marine microbes are still uncultured at this time and thus not amenable to genome sequencing. One attempt to fully characterize the microbial genetic diversity in the ocean has revealed a staggering array of microbial genes, many of whose functions are unknown at this time (Venter et al., 2004). It will be quite some time before this wealth of information is synthesized and fully interpreted, but benchmark studies such as that of Venter et al. (2004) provide a glimpse of the overall genetic capabilities present in a natural marine ecosystem, and indicate interesting features and possible strategies for more focused studies.

Metagenomic studies are beginning to provide a picture of ‘whole community metabolism’. The great potential that can be realized and some of the major obstacles that must be overcome in order to fully exploit genomic information have been recently outlined (DeLong, 2002; Nelson, 2003; Schloss and Handelsman, 2003; DeLong, 2004b; DeLong, 2004a). The latter issues include problems with data management, computational tools for extracting information, and the significant problem of how we integrate genomic and ecological information. Also, it is worth noting that evolution within a community occurs at the level of the species (however one might define microbial ‘species’) and not at the level of the community. There is a tendency within metagenomic studies to consider the total genomic information in a community as one giant entity from which one can select genes and processes at random. Community-level approaches
must not lose track of the fact that species are the functional ecological units.

In addition to these large metagenomic efforts, more modest attempts to assess the structure of natural microbial communities, and how they respond to environmental forcing, are being fully utilized at this time. There are two basic approaches to characterizing community structure used in these studies; single gene sequencing efforts (usually 16S or 18S rDNA) and fragment analysis of amplified DNA. Nucleic acids for both approaches are taken directly from environmental samples. These culture independent approaches provide at least three important functions. First, they provide practical approaches (i.e. realistic given constraints of most sequencing budgets) for assaying the spectrum of phylotype diversity in an ecosystem (a feature that is still very poorly known). Second, they enable comparative experimental studies of microbial communities by providing ‘snapshots’ of community composition at a given time and place, and thus a means of investigating how composition responds to changes in biotic and abiotic factors. Third, assuming that the specific method yields taxonomic information on the microbes present, these approaches provide a taxonomic/phylogenetic framework upon which to hang the enormous amount of genetic information produced by metagenomic studies of natural communities.

Cloning and sequencing of rRNA genes has become a fairly common method of assaying the diversity of natural assemblages of marine bacteria, archaea and protists. To date, this approach has been more extensively applied to prokaryotes (16S) for reasons noted above, but its application to eukaryotes (18S) has shown that there is still much to be learned about the diversity of all microbial taxa (Giovannoni and Rappe, 2000; López-Garcia et al., 2001; Moon-van der Staay et al., 2001). Cloning/sequencing has been used less frequently as a tool in comparative experimental studies of microbial assemblages because of the labor-intensive nature (and cost) of sequencing, but its application to natural assemblages of microbial eukaryotes has demonstrated dramatic changes in the composition of the assemblage during short-term (up to 3 day) incubations (Countway et al., 2005) and in response to perturbations to plankton food webs (Schnetzer et al., in prep.).

Amplification of nucleic acid fragments and fragment analysis of the resulting mixture of amplicons has been the more popular approach for comparative studies of microbial communities. A variety of methods, including terminal restriction fragment length polymorphisms (T-RFLP), denaturing gradient gel electrophoresis (DGGE), automated ribosomal intergenic spacer analysis (ARISA) and amplified ribosomal DNA restriction analysis (ARDRA) have been applied to prokaryote and eukaryote assemblages from a variety of ecosystems (Fisher and Triplett, 1999; Moeseneder et al., 1999; Watts et al., 2000; Diez et al., 2001; Scharfer et al., 2001; Muylaert et al., 2002; Blackwood et al., 2003; Cifuentes et al., 2003; Flaten et al., 2003; Gast et al., 2004; Countway et al., 2005). The power of these approaches comes from the speed with which an investigator can obtain community-level information on a microbial assemblage. This speed enables comparisons between different communities or treatments in an effort to examine how microbial community structure changes with environmental factors. While fragment analysis methods generally lack the resolution that sequence information affords, their relatively low cost and ability to assay much larger numbers of samples make them a useful experimental tool in marine microbial ecology for the foreseeable future.

Microbial diversity and molecular taxonomy

The emergence of vast amounts of sequence information from marine microbes will enable studies of community structure and function (as noted above), but also taxonomic descriptions. Sequences of genes conserved among microbial taxa provide potential ‘signatures’ for a molecular taxonomy that will be eminently more practical for ecological studies than traditional taxonomic schemes of culture (bacteria) or morphology (protists) (Blaxter, 2004). DNA sequences have been rapidly adopted by the scientific community as a proxy for bacterial and archaeal species descriptions because many of these taxa cannot presently be cultured and therefore are unavailable for physiological testing. The gold standard for this work has been and continues to be ribosomal DNA (usually 16S). The development of a molecular taxonomy has been met with less enthusiasm within the eukaryote community because of the long-standing morphology-based taxonomy in that field. However, rDNA gene sequences are now used regularly as additional taxonomic characters in phylogenetic analyses, and as putative taxonomic signatures for the design of techniques for the rapid identification and quantification of eukaryote species in natural assemblages (e.g. fluorescent in-situ hybridization, quantitative real-time PCR).
Development of a molecular taxonomy for microbial species is appropriate for a number of reasons. As noted above, many archaea and bacteria have not yet been brought into culture (a prerequisite for most biochemical testing) and offer little useful morphological detail. The use of sequence information is a logical approach to organize and categorize these taxa. For protists, taxonomic (morphology-based) expertise is waning, and thus a DNA taxonomy holds the promise of retaining vital taxonomic information on these important species. DNA sequence will also be able to identify morphologically-distinct life stages of protistan species as have been proposed for the ichthyotoxic dinoflagellate Pfiesteria piscicida (Burkholder et al., 1995), and distinguish cryptic species as have been noted for some planktonic foraminifera (de Vargas et al., 1999). Thus, a molecular taxonomy will be more inclusive of investigators from research communities that rely on microbial taxonomic information but who may not be taxonomic experts (e.g. ecologists, physiologists). Finally, genetic approaches for species identification will eventually be much more amenable to automation than traditional taxonomic schemes, greatly increasing the rate at which microbial community structure can be assessed in natural samples. Ideally, this approach should be highly analogous to, and coordinated with such programs as the Consortium for the Barcode of Life which has been initiated to provide a molecular taxonomy for cataloging metazoan life on the planet using cytochrome oxidase I as the gene of choice (Hebert et al., 2003).

The shortcomings of basing taxonomy or phylogeny on any single gene are well known. For example, lateral gene transfer can confuse relationships among taxa. This issue does not defeat the basic premise of defining a molecular taxonomy, but the inclusion of additional genes on which to found the taxonomy (or even some form of appraisal of the entire genome) may ultimately provide a more satisfactory scheme. The integration of information arising from a number of conserved genes should increase the robustness of any molecular taxonomy.

The availability of a molecular taxonomy for marine microorganisms will enable a range of research foci to move forward. A large database of DNA sequences will greatly improve the design of sequence-based approaches to identify and quantify target species in natural assemblages of microorganisms. A molecular taxonomy will also enable the development of macro/microarrays on which a large component of the microbial community can be assayed simultaneously. Such ‘phylochips’ will be instrumental for ecological studies of microorganisms and for rapidly establishing the presence or absence of species of interest (El Fantroussi et al., 2003; Urakawa et al., 2003).

It is worth noting that redefining microbial taxa based on sequence information will not resolve the difficult issue of the species concept for unicellular organisms (Cohan, 2002; Coleman, 2002). Most microbiologists appear to favor the ecological species concept, but the ecological niches of most microbes are poorly characterized, making this definition an unworkable taxonomy at present. Nevertheless, DNA-based taxonomy will allow biologists to correlate sequence identity with traditional taxonomic features (whether morphological or physiological) to design protocols that will provide the desired level of resolution for a particular goal or question.

**Expansion and exploitation of culture collections**

It can be expected that culture collections of marine microorganisms will see unprecedented growth in the coming years for scientific and practical reasons. A practical reason for establishing cultures is that culture is generally a prerequisite for genome sequencing. A more fundamental reason is that cultures remain the primary mechanism for studying the physiology and biogeochemical activity of microbial species. Mechanistic data cannot be obtained easily in nature, but the highly controlled environment of the laboratory allows careful, systematic examination of the effects of environmental variables on growth and behavior.

The recognition (based on culture-independent surveys of natural communities) that many marine microorganisms have not yet been cultured has spurred efforts to bring more microorganisms into culture. Novel approaches that involve cell sorting by flow cytometry and other high-throughput techniques, and culture in highly dilute media (a more normal environment than highly enriched media that have been used in the past) will greatly increase success in the isolation and cultivation of novel microbes (Connon and Giovannoni, 2002; Rappe and Giovannoni, 2003; Stevenson et al., 2004). Another major benefit to future culture attempts is the large amount of metagenomic data that will be gathered for various marine ecosystems. These efforts will help to establish the physiological abilities and constraints of microorganisms that are truly representative of the species that dominate natural
ecosystems. Metagenomics will undoubtedly reveal clues to help culture the ‘unculturable’ species.

In turn, cultures of novel marine microorganisms will be exploited by genomic approaches to understand their physiology and ecology at a level of sophistication that can only be imagined at this point. Physiological studies up to the present have employed relatively simple experimental methods directed towards understanding basic parameters of growth (e.g. elemental composition, nutrition, growth efficiency). These studies have sought to construct general principles for food web structure, energy flow and elemental cycling within and between microbial populations (Sterner and Elser, 2002). The application of genomic information will afford a wholly new level of understanding of the abilities, constraints and forces driving the physiology of individual taxa of microorganisms. Arrays will allow interrogation of microbes under a wide variety of conditions in such a way that not only will the outward responses of these species be observable but the underlying genetic and biochemical rules dictating those responses will become clear.

Science has long valued culture collections, but it is clear that governmental funding agencies now also grasp the value of supporting culture collections. Substantial funds are now available for supporting and expanding marine culture collections, and there is a strong expectation within the scientific community that voucher materials (including live/frozen cultures) will be deposited into public culture collections. Happily, this has also stimulated work on cryopreservation methods for species that have not previously survived these procedures. Significantly, private industry now recognizes the potential value contained within the vast amount of untapped microbial diversity present in nature, and is an eager participant in the culture of novel microbial taxa.

**Dissecting (and constructing) microbial communities**

The technological and computational advances in microbiology during the past few decades have been breathtaking, and the next few decades promise to be equally remarkable. We are entering an era where we can begin to ask and answer some of the most basic and enduring questions that microbiologists have pondered since microorganisms were first observed more than three centuries ago by Antonie van Leeuwenhoek. What is the true diversity of this great unseen community and what maintains its diversity? What are the evolutionary relationships among the microbial types that exist in nature? Are microbes globally distributed or do they exhibit endemism? How are ecological relationships between microbes (commensalisms, mutualisms, parasitisms, metabolic consortia) established and how are they maintained? Is the functional stability and/or resilience of a microbial community related to diversity of the community? What are the limits of microbial community function in the face of environmental change? How do communities adapt to such changes?

Answering these and other fundamental questions is finally becoming feasible with the array of new molecular approaches and technologies that are unfolding before us. For example, symbiotic relationships between macro- and microbial species have been documented for many years, but only recently are we beginning to understand the molecular basis for the relationships between certain microbial symbionts and their hosts through genomics, proteomics, bioengineering and biochemistry (Downie and Young, 2001; Lupp et al., 2003; Rolfe et al., 2003; Koropatnick et al., 2004; Ruby et al., 2005). Experimental approaches are beginning to hypothesize (and test!) the processes that influence the evolution of microbial species using model systems (Finkel and Kolter, 1999; Elena and Lenski, 2003). Moreover, model-driven biological discovery is rapidly becoming a mainstay of molecular biology as experimental data are subsumed into advances in computational biology, yielding mathematical models that predict the outcome or consequences of complex biological processes and providing testable hypotheses for further experimental studies. Such reductionist approaches focus on dissecting the complex interactions between species, or the effect of environmental conditions on evolution of species.

On the other hand, attempts to experimentally examine the relationship between a community’s biological diversity and its biogeochemical function, stability and resilience are focused on understanding the emergent properties of complex assemblages of species (McGrady-Steed et al., 1997; Naeem and Li, 1997). Much of the extant theory in this area is derived from mathematical (theoretical) ecology, with supporting evidence for theory obtained primarily from empirical observations and some experimental manipulation of (mostly) macrobenthic communities (Tilman, 1999; Worm and Duffy, 2003). However, microbes are exceptionally well-suited as
The measurement of microbial identity and abundance requires multiple processing steps to derive meaningful measurements of microbial taxa. A handful of instruments are presently capable of making measurements individually, collating biological, chemical and physical data in real or near-real time within the network, and eliciting a human response and/or activity of more sophisticated or costly instrumentation when appropriate (Estrin et al., 2004). These instruments and sensors are becoming a reality. Farther along in this time spectrum (but within the foreseeable future) will be the incorporation of ‘ecogenomic sensors’ of varied (and at this point, somewhat vague) design. These approaches will undoubtedly include micro/macaroarrays (‘phylochips’) to establish the presence and abundance of a vast array of common microbial types (El Fantroussi et al., 2003), but also approaches to establish the presence, diversity and activity of genes of specific function (Steward et al., 2004).

The benefit of this new generation of in-situ instrumentation to our basic understanding of microbial distributions, processes and the factors acting at smaller spatial and temporal scales than are presently addressed by most instrument deployments. In order to effectively monitor microbial taxa, and to obtain an ecological understanding of the factors governing their presence/abundance, it is necessary to study at these finer scales.

Thus, one can envision a large number of relatively inexpensive, autonomous, networked instruments and/or vehicles in the future that are capable of making measurements individually, collating biological, chemical and physical data in real or near-real time within the network, and eliciting a human response and/or activity of more sophisticated or costly instrumentation when appropriate (Estrin et al., 2004). These instruments and sensors are becoming a reality. Farther along in this time spectrum (but within the foreseeable future) will be the incorporation of ‘ecogenomic sensors’ of varied (and at this point, somewhat vague) design. These approaches will undoubtedly include micro/macaroarrays (‘phylochips’) to establish the presence and abundance of a vast array of common microbial types (El Fantroussi et al., 2003), but also approaches to establish the presence, diversity and activity of genes of specific function (Steward et al., 2004).

The benefit of this new generation of in-situ instrumentation to our basic understanding of microbial distributions, processes and the factors acting in terms of machine precision and energy use (e.g. they require careful sample handling and temperature control, even temperature cycling if PCR amplification of nucleic acids is involved). ‘Upstream’ concentration of samples via filtration or other methods is often required for detecting target taxa present at low abundance, and this requirement further increases energy demands. The next generation of biological sensors will require close collaborations between engineers and molecular biologists in order to make expectations a reality.

Nevertheless, significant progress in biological instrument development can be expected in the near future, not just within the area of biosensor development but also in the manner in which sensors will be deployed and the resulting information utilized. The present mindset for oceanographic instrumentation is to produce and deploy a few, very sophisticated instruments from which large extrapolations are made to represent the biology of whole ecosystems. While this approach is appropriate for some level of questions (e.g. large-scale climatic changes and their general effects on the biology of a ecosystem), most biological processes are driven by factors and forces that act at smaller spatial and temporal scales than are presently addressed by most instrument deployments.

A longstanding goal for ecological studies of aquatic microorganisms has been the desire to measure the abundances, distributions and activities of these species in-situ and in real time. This ability is necessary to monitor the status of those species that are important to ecosystem function and/or public health (e.g. human pathogens, harmful bloom forming algae), and to enable experimental studies that attempt to elucidate the factors that stimulate or promote the growth of these species. While this situation is becoming a reality for many physical and chemical features of the ocean (e.g., see Ocean Research Interactive Observatory Networks, ORION; http://www.orionprogram.org/), the development of rapid ‘biosensors’ for making in-situ measurements of microorganisms has significantly lagged behind.

Moored buoys, remotely operated vehicles and autonomous vehicles are essential (and common) tools of modern marine science. High-resolution measurements of ecologically-pertinent parameters such as temperature, pressure (depth), light, salinity, dissolved oxygen and various nutrient ions have begun to provide synoptic views of these variables in nature. Satellite imagery is also making great strides in the ability to provide large-scale depictions of surface features (some of biological importance) of the ocean (Dickey and Chang, 2001). In contrast, only a handful of instruments are presently capable of making meaningful measurements of microbial taxa in-situ and in near-real time (Scholin and al., 2000; Olson et al., 2003; Casper et al., 2004).

The impediments involved with developing these instruments are not trivial. Most biological assays require multiple processing steps to derive a measurement of microbial identity and abundance. These actions are demanding in terms of machine precision and energy use (e.g. they require careful sample handling and temperature control, even temperature cycling if PCR amplification of nucleic acids is involved). ‘Upstream’ concentration of samples via filtration or other methods is often required for detecting target taxa present at low abundance, and this requirement further increases energy demands. The next generation of biological sensors will require close collaborations between engineers and molecular biologists in order to make expectations a reality.

Nevertheless, significant progress in biological instrument development can be expected in the near future, not just within the area of biosensor development but also in the manner in which sensors will be deployed and the resulting information utilized. The present mindset for oceanographic instrumentation is to produce and deploy a few, very sophisticated instruments from which large extrapolations are made to represent the biology of whole ecosystems. While this approach is appropriate for some level of questions (e.g. large-scale climatic changes and their general effects on the biology of a ecosystem), most biological processes are driven by factors and forces that act at smaller spatial and temporal scales than are presently addressed by most instrument deployments. In order to effectively monitor microbial taxa, and to obtain an ecological understanding of the factors governing their presence/abundance, it is necessary to study at these finer scales.

Thus, one can envision a large number of relatively inexpensive, autonomous, networked instruments and/or vehicles in the future that are capable of making measurements individually, collating biological, chemical and physical data in real or near-real time within the network, and eliciting a human response and/or activity of more sophisticated or costly instrumentation when appropriate (Estrin et al., 2004). These instruments and sensors are becoming a reality. Farther along in this time spectrum (but within the foreseeable future) will be the incorporation of ‘ecogenomic sensors’ of varied (and at this point, somewhat vague) design. These approaches will undoubtedly include micro/macaroarrays (‘phylochips’) to establish the presence and abundance of a vast array of common microbial types (El Fantroussi et al., 2003), but also approaches to establish the presence, diversity and activity of genes of specific function (Steward et al., 2004).

The benefit of this new generation of in-situ instrumentation to our basic understanding of microbial distributions, processes and the factors
controlling them will be profound. The practical applications are tangible. Real-time monitoring of drinking water supplies, surveillance for the contamination of swimming beaches by potentially pathogenic microbes, and scrutiny of aquaculture and recreational/commercial fisheries for outbreaks of pathogens and toxin-producing algae are high priorities on federal and state agendas. There is a critical need for in-situ applications of genetic information, and this area will continue to grow and evolve rapidly as these instruments are designed, deployed and refined.

CONCLUSION

Marine microbial ecology has moved rapidly from the position of ‘intriguing footnote’ of limited importance in oceanography during the 1950s-60s to a pivotal aspect of ocean science and a focal point for cutting-edge biological and environmental research. Spurred on by a growing realization of the central role of microbes in ocean biology and biogeochemistry, and by remarkable breakthroughs in microbiology and molecular biology, we stand at the beginning of an outstanding growth period in this field. Moreover, the ease with which many of these species can be manipulated physically and genetically makes them ideal candidates for research on numerous aspects of basic biology. The greatest difficulty may be that of individual investigators keeping pace with the multidisciplinary nature of the field and the constantly changing landscape of knowledge and techniques. Undoubtedly, it will be an exciting and rewarding challenge, and an area that will witness remarkable advances during the next decade and beyond.

ACKNOWLEDGEMENTS

The author is helpful to grateful comments on the manuscript provided by Peter Countway, Ilana Gilg, Adriane Jones, Stephanie Moorthi, Astrid Schnetzer and Beth Stauffer. Production of this manuscript was supported in part by National Science Foundation grants MCB-00084231, OPP-0152437, The Center for Embedded Networked Sensing (CENS) under the NSF Cooperative Agreement CCR-0120778, National Oceanic and Atmospheric Administration grant NA160P2790, and Environmental Protection Agency grant RD-83170501.

REFERENCES


