

# Patterns and Prediction in Microbial Oceanography

BY JOHN J. CULLEN, W. FORD DOOLITTLE, SIMON A. LEVIN, AND WILLIAM K.W. LI

Microbes are integral to the structure and functioning of marine ecosystems, and to the chemistry of the ocean and its interaction with the atmosphere. They mediate the chemical transformations that over geological time have determined the composition of our atmosphere (Fennel et al., 2005) and the balance of major nutrients in the sea (Arrigo, 2005). These nutrients ultimately determine how much and what kinds of marine life can be supported in surface waters, the amount of primary production that can be transferred to higher trophic levels, and how much carbon is stored in the deep ocean (Falkowski et al., 1998). Microbes in the ocean are thus directly and indirectly sensitive to, and part of, the ocean's response to global change. The manifestations of global change include not only the changes in ocean temperature, circulation, pH, and nutrient availability linked to increasing greenhouse gases, but also the profound alterations of the marine environment associated with a growing and rapidly developing human population—habitat destruction, coastal eutrophication, pollution (GESAMP, 2001), and the elimination of larger predators by fishing, leading to fundamental alterations in food webs (Pauly et al., 1998; Myers et al., 2007).

Our challenge as scientists is to

describe and understand the environmental control of marine microbial communities so that we can predict the influences of global change on the biology and chemistry of the ocean as it interacts with and influences the atmosphere and climate. The central questions are clear, and they share a common theme: How will marine microbes, and the ecosystems they support, respond to environmental change, and what are the consequences for ecosystems and the biosphere? Broadly, we need to describe and understand the structure (species composition, size distributions, and flows of energy and materials) and functioning (nature and efficiency of chemical transformations) of microbial communities, recognizing that structure and functioning are intimately linked.

## FUNDAMENTAL PATTERNS IN MARINE SYSTEMS

Much progress has been made in describing regularities of marine ecosystems that are robust in their central tendencies, thereby reflecting fundamental structuring processes in the ocean. Well-studied relationships include the Redfield ratios (Redfield, 1958), the Sheldon size spectrum (Sheldon et al., 1972; Platt and Denman, 1978), and patterns of phytoplankton abundance in relation to oceanographic influences

(Sverdrup, 1955; Margalef et al., 1979; Longhurst, 2007) (Figure 1). These relationships are by no means exact; they vary in space and time with physical and chemical conditions, and are thus signatures of those conditions and the ecological interactions that are driven by them. Ramón Margalef showed directly the results of such physical-chemical-ecological interactions with his influential “mandala” (Margalef et al., 1979) that describes the selection of phytoplankton communities by regimes of turbulence and nutrients (Figure 1D). The relationships described by Margalef's mandala are general, but very well suited to guiding more detailed modeling for ecological prediction (Cullen, in press).

In a landmark paper, Redfield (1958)

---

JOHN J. CULLEN (*John.Cullen@Dal.ca*) is Professor, Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada. W. FORD DOOLITTLE is Professor, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada. SIMON A. LEVIN is Moffett Professor of Biology, Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. WILLIAM K.W. LI is Research Scientist, Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada.

(see also Williams, 2006) described the most famous, and perhaps most controversial, macroscopic features of marine ecosystems. Redfield compared the average N:P ratio of plankton (16:1 by atoms) and that of the dissolved nutrients, nitrate, and phosphate in deep oceanic waters (15:1), and the regularity of these features across broad regions of the ocean. He intuited that this was no accident, and that it was not just a case of the plankton responding to physically determined environmental variation. Rather, he hypothesized that the plankton were regulating the stoichiometry (here, meaning the N:P ratio) of the physical environment at just the conditions they needed. Similar arguments applied to the regularities in other element ratios, for example, C:N. The question remained, though, as to how this might arise, because natural selection operates for the benefit of the individual genome, and not to shape ecosystem properties.

The picture is not, of course, as simple as first presented. There is substantial variation among phytoplankton in their N:P ratios, and 16:1 is an average bracketed by extreme values that span more than an order of magnitude. There is, similarly, a good deal of variation in oceanic conditions, enough so that on his laboratory Web site, University of Hawaii oceanographer David M. Karl avers that, “The Redfield ocean is on its last legs as a construct for biogeochemical modeling.” Nonetheless, the Redfield patterns are a valuable approximation and point of departure. Klausmeier and colleagues (2004a, 2004b) build on classical chemostat models of the ocean, incorporating growth according to a standard law (Droop, 1974) and Liebig’s Law of

the Minimum (which states simply that growth will be limited by whichever essential nutrient is in shortest supply), to show that, in general, the system’s dynamics will reach an equilibrium in which a single factor is limiting. When this dynamic is embedded into an evolutionary framework, however, in which organisms must expend their available carbon on making nitrogen-rich proteins (the primary machines responsible for uptake of nutrients) or phosphorus-rich ribosomes (which play a key role in growth), the system evolves to co-limitation, in which neither nutrient is in excess. The reason for this is intuitively simple—if, for example, nitrogen alone were limiting, a mutant that relied less on nitrogen and made better use of the available phosphorus could proliferate. Most importantly, the optimal allocation to proteins and ribosomes is determined primarily by the stoichiometry of proteins and ribosomes, and is largely inde-

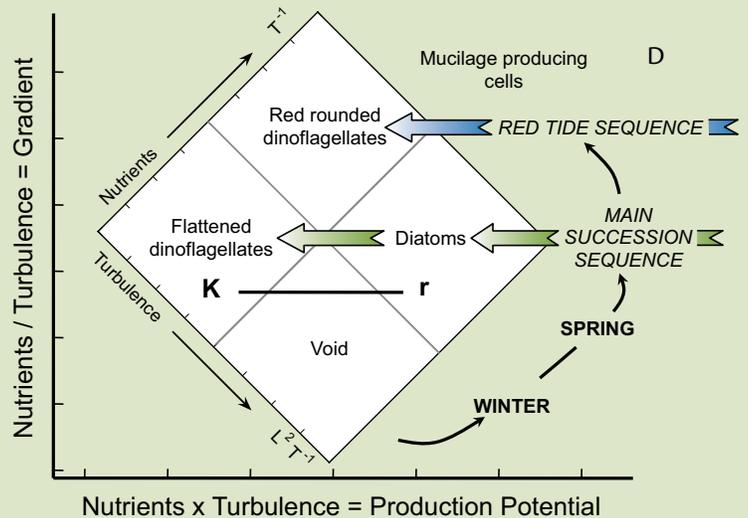
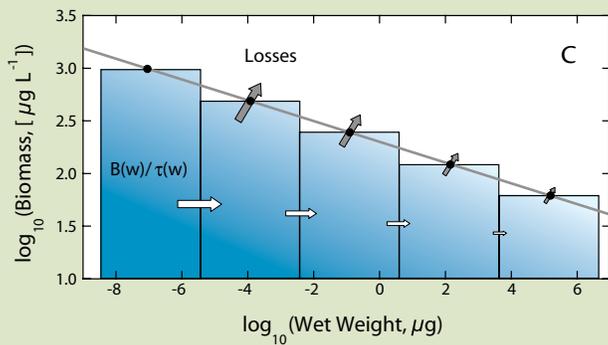
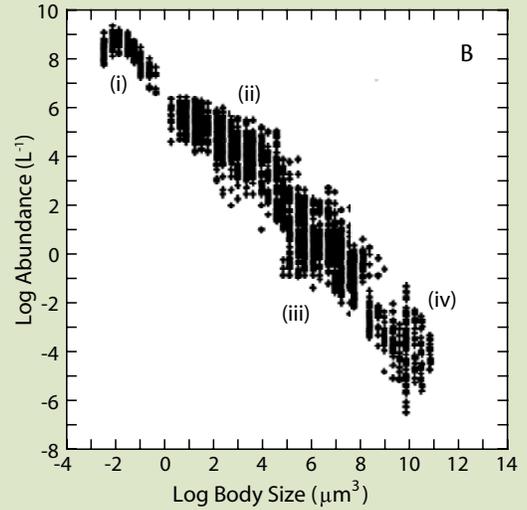
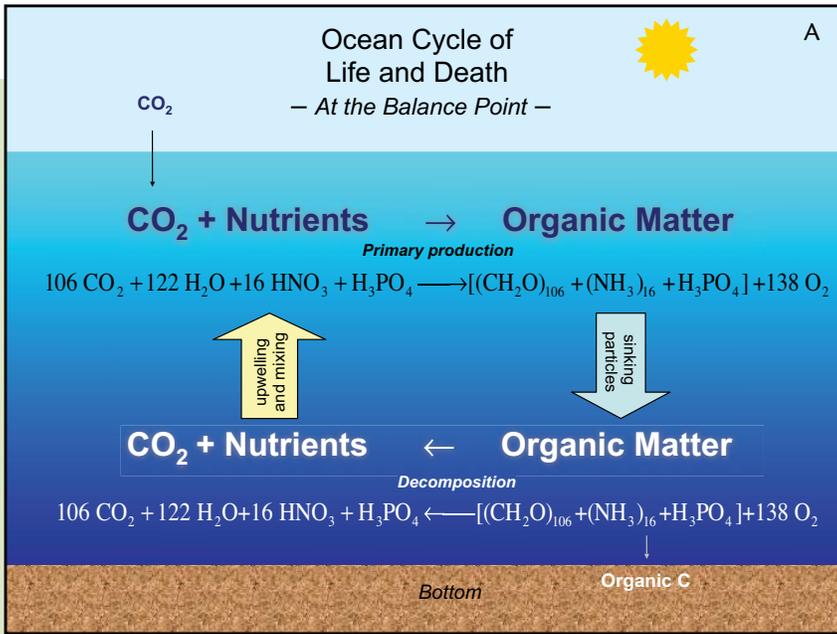
pendent of environmental conditions. Thus, at first blush, there is a justification for Redfield’s interpretation.

On closer examination, however, some gaps remain. This simplistic approach predicts an N:P ratio that overestimates Redfield by a factor of three, primarily

because it emphasizes equilibrium conditions in which competition selects for efficient acquisition of nutrients. Nor does this approach explain the coexistence of multiple types. Enter the great limnologist, G. Evelyn Hutchinson, who also was intrigued by a similar paradox (Hutchinson, 1961), and invoked environmental variation as part of the explanation of coexistence. The work already described assumes a (marine) world in balance, but Hutchinson reminds us that local disturbances caused by everything from turbulence to marauding whales (highlighted by Smetacek and Nicol, 2005) complicate the simple picture. This knowledge should change our view of marine communities fundamentally: if disturbance is superimposed upon the simplistic dynamic of species interactions, removing the competitive species that would otherwise dominate at a low-diversity equilibrium, then there will be selection for rapid growth under

...until recently, there was no basis for describing in a mechanistic way the ecological and biogeochemical links between processes on the molecular scale all the way to global change.

some conditions. Indeed, if within the Klausmeier et al. (2004a, 2004b) framework one optimizes allocation for rapid growth, selection favors lower N:P ratios, about half of the Redfield prediction. The equilibrium view predicts a ratio too high, and the growth-rate approach



predicts one too low. This suggests that, in a spatio-temporal dynamic in which localized disturbances initiate temporal successions of species of increasing competitive ability but decreasing growth rates, coexistence will be possible and the model will predict the full range of empirically observed N:P ratios.

On the regional scale, such as in

the North Pacific subtropical gyre, the chemical composition of phytoplankton (and ultimately the N:P of the pools of dissolved and particulate organic matter in the surface layer) responds to relative supplies of N and P from below—that is, whichever runs out first. N<sub>2</sub>-fixing cyanobacteria, with inherently high N:P, are favored in oligotrophic waters with low

N:P supply; their growth and subsequent influences on the chemical composition of deep water (increasing its N:P) lead to a hypothesized N<sub>2</sub>-fixation cycle between N-controlled and P-controlled systems (Karl, 2002) that can be explained with physiological principles that are by no means inconsistent with the broader frameworks described by Klausmeier et

Figure 1. Fundamental regularities in ocean biology and chemistry reflect the structure and functioning of marine ecosystems and biogeochemical cycles, but they vary in space and time with physical and chemical conditions—and it is this variability that determines the richness of the ocean's responses to environmental change. (A) Marine microbes, primarily phytoplankton in surface layers of the ocean, convert dissolved gases and inorganic nutrients into food. Other microbes break it back down again, governing the cycle of life and death in the sea. By mediating these fundamentally important flows of energy and conversions of matter, marine microbes serve as the foundation of the ocean's food webs and the engines of its biogeochemical cycles. When considered globally and in long-term balance, this cycle of life and death in the sea follows Redfield stoichiometry with  $\text{CO}_2$ , nitrate, and phosphate being assimilated into organic matter—phytoplankton—and ultimately regenerated in the deep sea with an atomic ratio of close to 106:16:1. In steady state, the upward transport of dissolved inorganic carbon matches the downward flux of carbon in organic matter, so there is no net downward transport of carbon associated with this internal cycling (Eppley and Peterson, 1979). The cycle maintains the biologically mediated gradient of  $\text{CO}_2$  that was established by depletion of nutrients in the surface layer and decomposition of the resultant organic matter at depth. Environmental variability ensures that steady state does not exist, however, and the resultant changes in nutrient distributions and departures from Redfield have profound influences on the nitrogen inventory of the ocean, on carbon fluxes, and on food web interactions in the sea. (B) Aquatic systems show remarkable constancy in size spectra, the abundance or amount of biomass in successive logarithmic size bins. This spectrum describes the size distribution of the abundance of planktonic organisms ranging from bacteria (i), to small phytoplankton (ii), to large phytoplankton (iii), and to mesozooplankton (iv), compiled from oceanic and nearshore waters, reservoirs, and high mountain lakes (from Figure 5 in Rodríguez, 1994). The relationship has a slope that closely approaches the value of -1, which is consistent with theory proposed by Sheldon et al. (1972) and subsequently studied by many others. (C) Assuming that energy and biomass flow to larger organisms from smaller ones, the steepness of the slope between size bin ( $w$ ) and the next larger ( $w+1$ ) in this biomass spectrum is determined by its biomass,  $B(w)$ , divided by its turnover time,  $\tau(w)$ . This consumption-related flux between compartments reflects losses due to incomplete assimilation of ingested material and respiration. A more steeply negative slope can imply less-efficient energy flow to higher trophic levels. A bulge in the spectrum implies an accumulation of energy/biomass following a pulse in production (e.g., a spring bloom). This bulge can travel to larger size classes via consumer links, and is attenuated/diffused according to the properties of each consumer size class. *From recent work of author Levin and Charles A. Stock and Thomas M. Powell, both University of California, Berkeley; paper submitted* (D) Margalef's mandala (redrawn from Margalef et al., 1979) by Cullen (in press) describes how the community structure of phytoplankton is characterized by functional morphology as it relates to turbulence and nutrients—environmental factors that have a very strong influence on survival of phytoplankton and that are expected to vary with global warming. This model provides no quantitative predictions, but it is generally consistent with observation, experimentation, and theory (Kiørboe, 1993), so it is useful as a guide in predictive modeling.

al. (2004a, 2004b) and Arrigo (2005).

Stoichiometric ratios are just one example of macroscopic patterns that exhibit regularities in marine systems, and that emerge from interactions mediated at lower scales of organization. Similar regularities are observed in the spatio-temporal patterns of patchiness for which unique signatures are captured

in spectral plots of the distributions of zooplankton, phytoplankton, and passive markers of oceanic conditions. The existence of patterns in the distribution of planktonic organisms has been a central focus of research for a half-century or more, in part because of the phenomenon of red tides and other noxious organisms that cause extensive mortal-

ity to wildlife, both marine and coastal. Phytoplankton are patchy in their distributions, but zooplankton are even patchier because they actively aggregate. Similarly, graphs of the relative abundances of all species assume canonical shapes for marine systems, just as they do for terrestrial systems, statistical regularities that summarize patterns of biodiversity. Most intriguing, perhaps, for marine systems, is the consistency in the power-law distributions that describe the size spectra, from the smallest microorganisms to the largest fish in the oceans (Figure 1B, C). Sheldon et al. (1972) showed that the biomass spectrum in aquatic systems exhibits remarkable constancy in biomass in successive logarithmic size bins. Numerous explanations have been advanced for this regularity; although a universally satisfactory general explanation seems still lacking, these theoretical investigations have led to better understanding of the forces structuring marine ecosystems.

When sufficiently well developed, the hypotheses that explain macroscopic regularities can be used to describe and predict how ocean systems will respond to environmental change. An example is the Redfield-based description of the global balance between  $\text{N}_2$  fixation and the microbially mediated loss of fixed nitrogen, as illustrated in Figure 2 and discussed by Arrigo (2005). This theoretical framework links environmentally determined controls on the growth rates of phytoplankton (i.e., replenishment of surface waters through upwelling or vertical mixing), phytoplankton N:P stoichiometry, and ultimately the N:P of the deep ocean. It can be exploited to hypothesize effects of another alteration

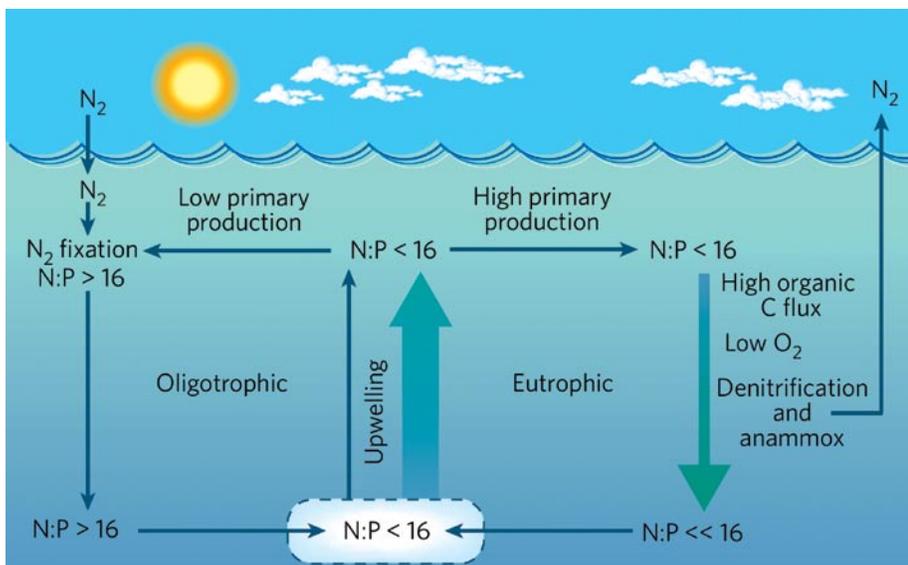


Figure 2. Departures from Redfield stoichiometry are highlighted in this description of microbially mediated controls on the ocean inventory of biologically reactive nitrogen (from Arrigo, 2005). The example is for the contemporary ocean, which has a mean N:P slightly less than 16. Waters enriched by upwelling or mixing support high rates of production of organic matter with relatively low N:P. Sinking of this matter and its subsequent degradation leads to oxygen depletion, loss of fixed nitrogen from the ocean through denitrification and anaerobic ammonium oxidation (anammox) reactions, with further reduction of deep ocean N:P. These processes are countered in oligotrophic waters, where stratification and nitrogen depletion favor  $N_2$  fixation and the production of organic matter with high N:P. The trends in the stoichiometry of primary production are consistent with analyses of microbial growth and chemical composition in an evolutionary framework (Klausmeier et al., 2004a, 2004b), reinforcing the view that the ecology and evolution of marine microbes are integral to biogeochemical cycling in the sea. *Reprinted by permission from Macmillan Publishers Ltd: Nature, (Arrigo, 2005), copyright 2005*

of the marine N cycle, modulation of  $N_2$  fixation in oligotrophic gyres by supplies of iron from atmospheric deposition, or, conceivably, intentional fertilization. The framework suggests that any such changes big enough to make a difference to global ocean productivity and carbon flux would have far-ranging influences on the N:P stoichiometry of the global ocean and thus microbial communities in the surface layer and the deep sea, with possible effects on both food webs and denitrification, respectively.

Hypotheses like these are generally well matched to broad questions related to global change, questions whose answers would be predictions of the

nature of change, often with a quantitative and testable estimate of its magnitude. The fundamental rules underlying macroscopic descriptors do not directly yield detailed simulations of ecosystem response. Rather, they guide the development of mathematical relations that, when applied in numerical models, reproduce the features. Regardless, to develop and validate macroscopic descriptors, it is necessary to measure relevant characteristics of microbial communities, along with the physical, chemical, and biological factors that influence them, on scales from those at which the patterns are determined to those on which they are expressed.

## DESCRIBING PATTERNS ON A RANGE OF SCALES

The search for macroscopic regularities can be considered an exercise of increasing the scale of perception. By scale, we mean the distance or time over which a significant change in the quantity of interest can be resolved. Therefore, the task is to extend the scope of space and time in our observations (Li, 2007). This is a familiar concept in ordinary experience because we easily grasp the notion of day-to-day fluctuations of weather occurring over the backdrop of a regional climate.

In human terms, atmospheric climate change is a slow phenomenon. Scientists discern it from analysis of much faster local weather events. Atmospheric conditions are monitored continuously or at regular, short intervals over a long time period at many locations. In the marine environment, physical oceanographers derive hydrospheric climatologies in a similar fashion using a grid of observing stations that provide continuous, long-term measurements of, say, temperature and salinity. To document the effects of climate change on marine microbes, biological oceanographers face the same task of constructing climatologies in order to discern anomalies. Unfortunately, many biotic properties, such as distributions of species and rates of production, are much more difficult to monitor accurately than physical properties. Nevertheless, for both temperature and microbes, the norms (i.e., climatology) must first be established before the changes (i.e., anomalies) can be detected. Interdisciplinary environmental surveys need to be repeated at regular intervals in order to yield infor-

mation appropriate for the assessment of climate change.

To describe linkages between climate and microbial community structure and function, there must be appropriate data to encompass microbial generation times of a day or so to ecosystem response times of decades or more. The relationships among physics, chemistry, and biology in small water parcels need to be scaled to ocean basins. A decade or two ago, this was barely conceivable. With new technologies for automated measurements from remote platforms such as satellites, moorings, and vertically profiling ocean gliders, multi-scale interdisciplinary sampling of the ocean is now a reality (Babin et al., 2005; Dickey et al., 2006).

Marine microbes can in principle be detected in a very large number of ways, depending on which of their cellular or physiological features is targeted. However, a requirement of climate monitoring is to assure that the same quantity is always measured. The value of a time series increases with its length, but only when it is internally consistent. Clearly, there is considerable burden in the initial selection of an appropriate variable because this sets an operational definition for what is being recorded. It is therefore not surprising that oceanographic monitoring often relies on high-level taxonomic indicators (such as the presence of chlorophyll *a*, indicating phytoplankton), which are generally more easily measured. In some cases, different variables may be merged retrospectively to construct a single extended time series for a particular biological entity (Raitsos et al., 2005). But even if it were not possible to do so, proxy indica-

tors can still provide a good assessment of climate-change impact on ecosystems. This is because data that identify broad systematic trends across diverse biotic groups and geographic regions provide a coherent “fingerprint” that can be assessed in probabilistic terms (Parmesan and Yohe, 2003). Inevitably, as technology advances, expensive analytical methods currently used only in exploratory surveys will be engineered into cost-effective applications for routine environmental monitoring. So, our job as scientists is to hasten efforts to document and understand past and present conditions in order to provide the foundations for future advances.

#### CASE STUDY:

##### MULTI-SCALE COARSE-GRAINING

Here, we present a case study for discerning long-term change from an extended time series of short observations. We consider one particular group of marine microbes, the so-called picoeukaryotic phytoplankton. This group is a diverse mix of different primary producers, excluding bacteria, that share the common physical characteristic of being very small (less than about two micrometers in diameter). Recent developments in molecular biology enable single taxonomic units to be differentiated from others in this group (Not et al., 2004; Zhu et al., 2005; Countway and Caron, 2006), but only limited data sets currently exist. However, the group as a whole is easily enumerated by flow cytometry, simply on the basis of small size and low photopigment content, discounting those cells with bacterial characteristics. Thus, time series of total picoeukaryotes can be used to construct

climatologies suitable for detection of interannual changes.

In Bedford Basin, Canada (Li et al., 2006), a 14-year record of weekly observations shows a recurring annual cycle of picoeukaryote abundance with a general trend of increasing abundance (Figure 3A). It is evident that the cycle is not exactly the same every year. For example, the annual minimum and maximum of cell abundance change slightly from year to year. These are the kinds of small shifts we need to detect. A baseline period for this series can be arbitrarily defined as 1993 to 2000 to provide a norm for comparison with subsequent observations. Over this period, each week in the 52-week cycle will have eight observations, one from each year, and the average of these constitutes the norm for that week. Fifty-two such weekly norms constitute the baseline climatology for picoeukaryote abundance (Figure 3B). The deviation of each weekly observation from the norm is termed an anomaly, and the average of all 52 anomalies in a year constitutes the annual anomaly. It is from the time series of annual anomalies that a deseasoned trend of interannual change may be sought (Figure 3C). Positive anomalies of picoeukaryotes are increasingly evident in the early twenty-first century relative to the preceding normative period (Figure 3C). The 14-year record of cyclical changes in plankton invites questions about the possible effects of climate change: Have abundances been increasing for a long time? Are the spring increases coming earlier? In botanical studies, old photographs and herbarium specimens may sometimes be scrutinized for evidence

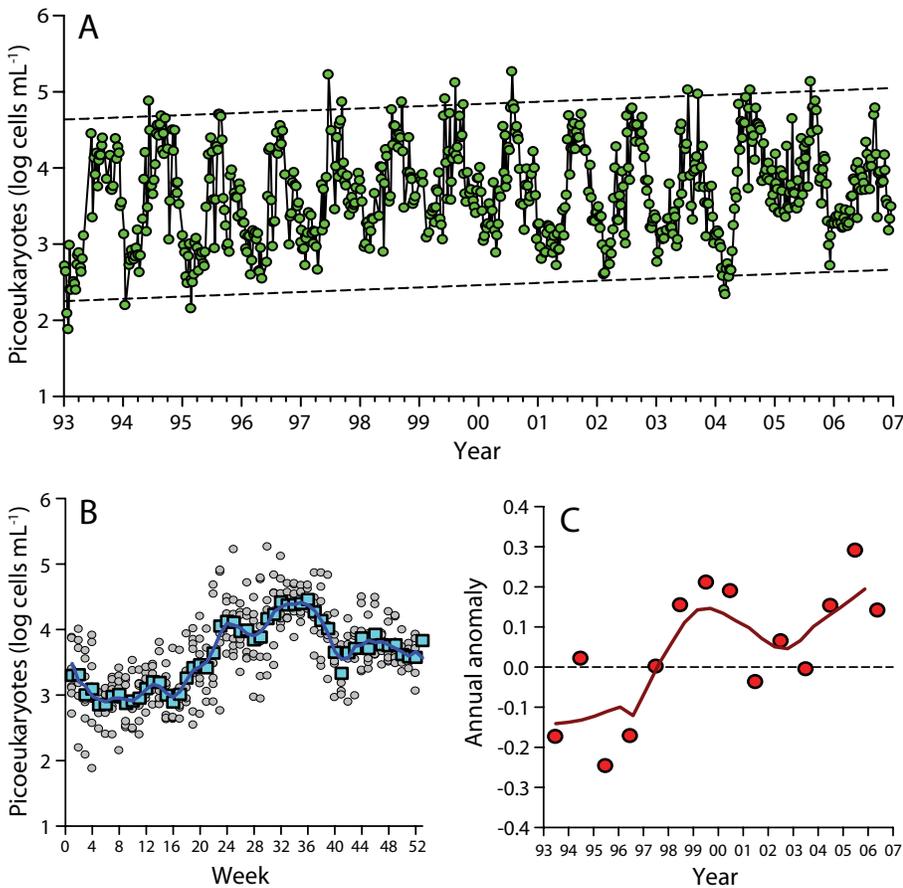


Figure 3. Abundance of picoeukaryotic phytoplankton in the Bedford Basin from 1993 to 2006. (A) Time series of weekly observations (green circles), showing 95% prediction lines from linear regression. (B) Climatological norms based on data from 1993–2000. Symbols: weekly observations (grey circles), weekly norms (blue squares), three-point running average of norms (blue line). (C) Annual anomalies (red circles), with trend indicated by locally weighted regression (red line).

of climate-related changes in the timing of flowering in plants (Miller-Rushing et al., 2006), but retrospective analyses of microbial plankton are largely unknown because documentable resources do not exist for them.

The year-to-year change in the annual average value of an ocean variable is almost always much smaller than the within-year seasonal variability. This is true of temperature, and also true of microbial abundance. This means that multiyear trends cannot be predicted solely from knowing how microbes

respond to short-term environmental change. From our case study, within any year, weekly picoeukaryote abundance is strongly correlated with weekly temperature (Figure 4A), but not so with weekly nitrate concentration (Figure 4B). However, these relationships do not predict what happens over many years. The multiyear pattern is only seen when weekly observations are averaged over the whole year. On an annual basis, by and large, picoeukaryotes increase and decrease coherently with nitrate (Figure 4D), but not with temperature

(Figure 4C). In the Bedford Basin, it seems that a secular rise in nitrate is imposing an upward tilt to the recurring temperature-linked annual cycle of picoeukaryotes. Thus, environmental change can lead to different biological outcomes, depending on the time scale. This example illustrates why relatively short-term trends in climate-related variability should be extrapolated to longer-term global change only with caution and due regard for mechanistic links (Behrenfeld et al., 2006).

This case study also demonstrates that statistical robustness is strengthened by having many observations in a year, and by having many years of observation. Current technologies for remote observation of marine microbes, such as submersible robotic flow cytometers (Dubelaar et al., 1998; Olson et al., 2003), are equipped to collect many observations during remote deployments, but they cannot be left unattended for years and decades. Regardless, there will always be a need to integrate wet microbial oceanography with remote Earth observation systems. It is a strategy to balance ecological complexity and simplicity.

## GENOMIC DIVERSITY IN MARINE MICROBES

Just as it was once hopelessly unrealistic to think about describing the distributions of marine microbes in relation to their physical and chemical environment on scales from centimeters to thousands of kilometers, until recently it was not possible to imagine characterizing these microbes except in the most general functional terms. Although, in total, a few thousand strains of marine microal-

gae, other microbial eukaryotes, bacteria, and archaea have been studied to some extent in culture, and more than 100 will have had their genomes sequenced by the end of this year, most marine microbes (perhaps more than 99%) remain uncultured, seem unculturable, and are certainly uncharacterized—taxonomically, physiologically, and genetically.

Molecular, genomic, and now metagenomic methods have changed all this. Phylogenetic marker genes, principally those for small subunit ribosomal RNA (SSU rRNA), permit facile classification of cultured cells. Most importantly, amplification of such marker genes directly from environmental DNA samples using the polymerase chain reaction and their characterization (often by determining DNA sequence) allows assessments of phylogenetic affinities, relative abundances, and diversity that are not subject to the biases of culturability (Pace, 1997). There are now in the literature more molecular surveys of marine microbiota than “traditional” (culture-dependent or spectroscopic and biochemical) studies.

But, there is in this diversity an embarrassing richness, especially for prokaryotes. Not only are there far more and far more different kinds of microbes out there (on the basis of SSU rRNA species profiles) than we would have ever guessed, there is much more genomic diversity hidden within the genomes of what, by all accounts, should be members of the same species than we can yet possibly explain. Few bacterial species (as traditionally circumscribed and named) comprise a single, homogenous population. Most appear to be assemblages of hundreds of strains, differing among

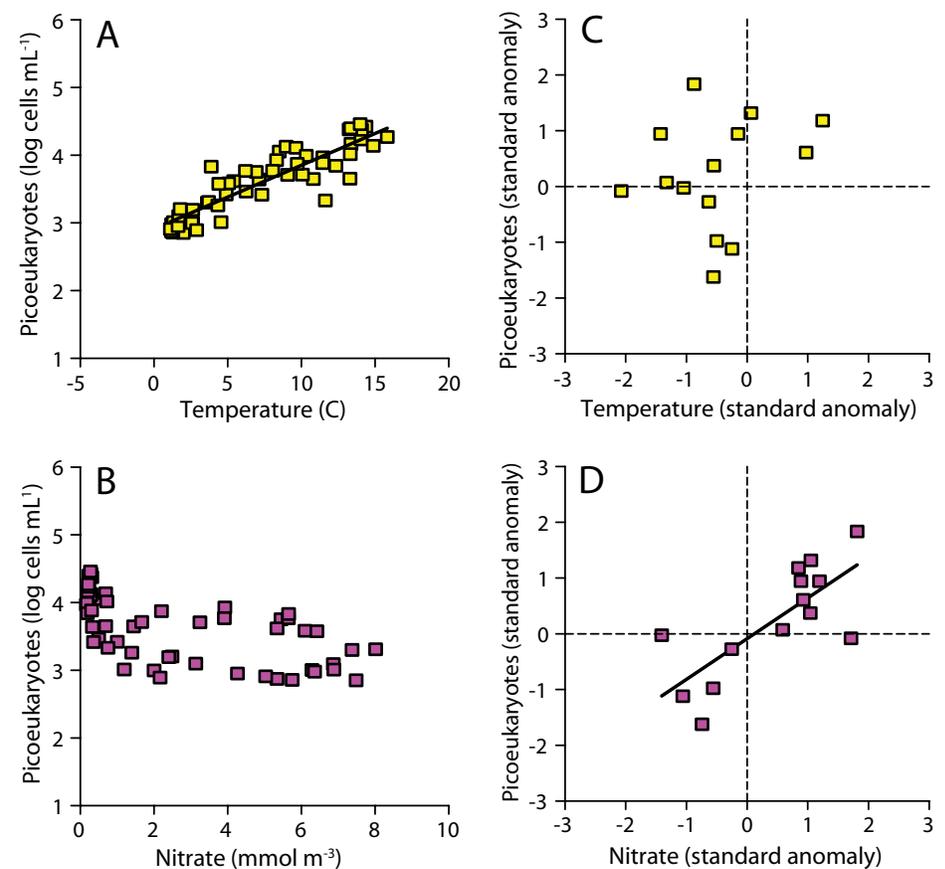


Figure 4. Importance of temperature but not nitrate on the abundance of picoeukaryotic phytoplankton in the Bedford Basin at an intra-annual time scale, indicated by weekly norms (A, B); conversely, importance of nitrate but not temperature at an interannual time scale, indicated by standard annual anomalies (C, D).

themselves by as much as 30% in gene content—more variation in hereditary information than might be found in all the vertebrates. Many of these closely related strains may, and probably do, have different ecological niches, thus comprising different “ecotypes.” Each strain of a species will have, in addition to a “core” of genes shared with all other strains of that species, auxiliary genes present only in some (or no) others. This core, plus all the auxiliary genes found in all strains collectively, constitute a virtual “pangenome,” many times larger than any real genome (Fraser-Liggett, 2005).

The commendable efforts of the

Gordon and Betty Moore Foundation in genomics have largely focused on phylogenetic breadth rather than within-species or within-genus level diversity. But, there are now some dozen or more sequenced genomes for representatives of the ocean’s most abundant phytoplankton, cyanobacteria of the genera *Prochlorococcus* and *Synechococcus*. These genome sequences give us confidence that gene exchange and within-species gene content variability will be key features in the evolution and ecology of the sea, as they are already known to be for the better-studied, disease-causing microbes (Figure 5). Among

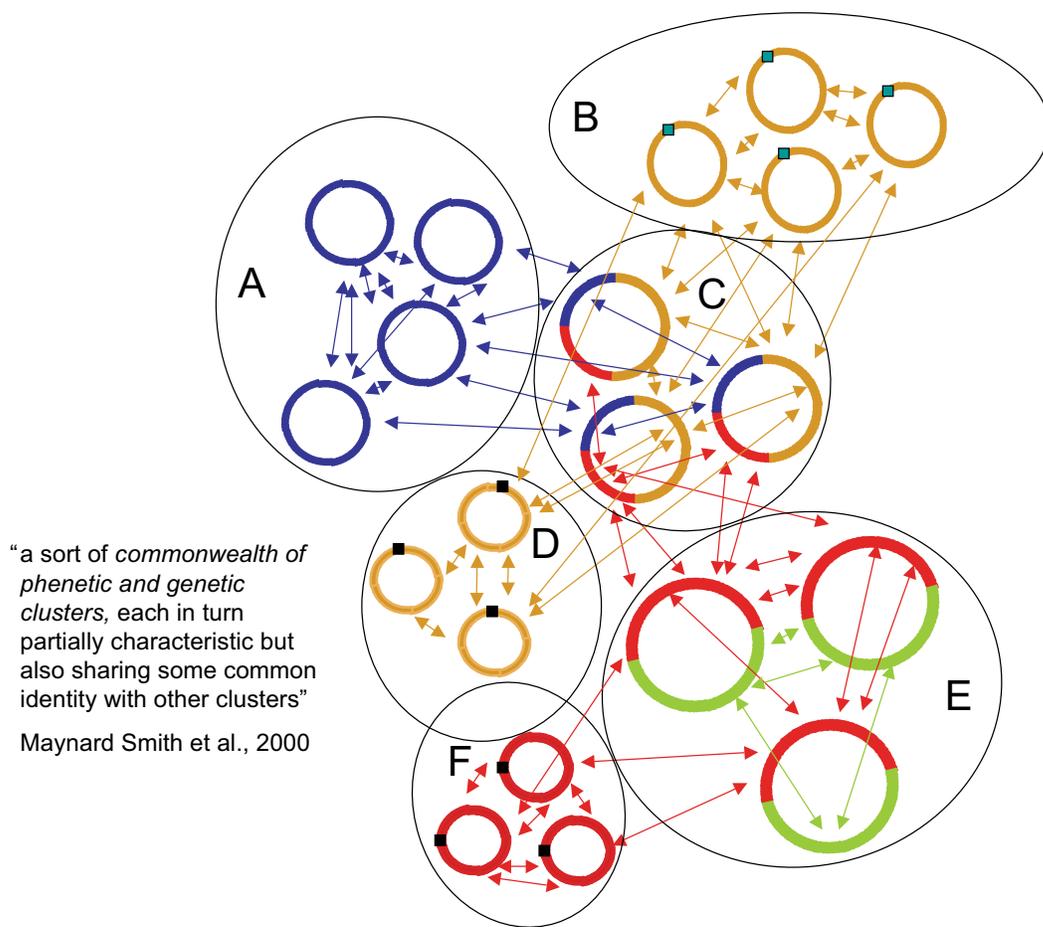


Figure 5. Bacterial “species” and gene exchange. Maynard Smith et al. (2000) describe the bacterial genus *Neisseria* (which includes the causative agent of gonorrhoea) as a commonwealth, and we indeed envision the relationships between bacterial assemblages that might or might not be called “species” as commonwealth-like. In this figure, similarly colored circles represent bacterial genomes that are sufficiently similar in DNA sequence that recombination (exchange of genes or gene segments) can readily occur between them (indicated by arrows). Circles with multiple colors represent mosaic genomes, different parts of which can exchange with similarly colored regions of other genomes. The enclosed regions A–F are populations, consisting of individuals that interact with each other most frequently, ecologically and genetically. The blue and black boxes represent genes laterally transferred from more remote genomes (outside the “commonwealth”). Such genes may often convey abilities, such as growth on a novel substrate or resistance to antibiotics, that define a population’s ecological niche. Because D and F have the same laterally transferred genes, they might exhibit overlapping niches, and yet be unable to exchange information by recombination, while E and F, though recombining, occupy different niches. In general, within-population gene exchange is more frequent than between-population exchange, but there is no reason not to expect that all possible combinations and extents of within- and between-population exchange and degrees of ecological separation will be observed among the ocean’s microbes. The quest for some uniform species concept can only confuse our understanding of such situations.

*Prochlorococcus* strains that might be considered a single species (> 97% SSU rRNA sequence identity), there are many ecotypes, differing in abundances along gradients of light, temperature, and nutrients, and accordingly in geographic distribution (Coleman et al., 2006;

Johnson et al., 2006). Genome comparisons confirm that such strains also differ very substantially in gene content and in physiologically relevant encoded traits, such as pigmentation, and nitrogen and phosphorus utilization. And indeed, at least one physiological attribute,

the optimum temperature for growth determined in the laboratory, correlates with distributions of ecotypes in nature (Johnson et al., 2006).

Strain-specific, environmentally relevant genes are often found clustered together in laterally transferable

“genomic islands” (Coleman et al., 2006). Recent calculations show that a 1% advantage due to a laterally transferred genomic island bearing genes for phosphate (or nutrient uptake) could effect replacement of one *Prochlorococcus* strain by another in an ocean basin in just a few decades, a time period comparable to that of the “observed domain shift in the North Pacific Ocean gyre from a nitrogen- to a [phosphorus]-controlled state” late in the last century (Martiny et al., 2006).

Certainly, for some genes present in only some species’ pangenomes (such as those of phosphate acquisition in *Prochlorococcus*), we will be able to find ecological correlates. Whether others are just “passing through” remains unclear, and there have been calls for a “neutral theory of gene content” to explain persistent genetic diversity among assemblages of what should be ecotypically very similar cells. For example, among *Vibrio splendidus* isolates with > 99% SSU rRNA identity from a single coastal site in Massachusetts, there may be “a thousand distinct genotypes, each occurring at extremely low environmental concentrations (on average less than one cell per milliliter)” —and some differing from each other by nearly 1 Mb (a thousand genes’ worth) of DNA (Thompson et al., 2005)! It strains credulity that there are that many distinct and tiny *V. splendidus* niches in a small patch of open water: some of this variation must be of no consequence. But, some is consequential, and mapping microbial genome content to ecology and location will be a much more challenging enterprise than it looked to be when we thought that all we needed to

do was identify what species were found where and rationalize this in terms of the information in the sequenced genome of the species’ “type strain.” Single markers such as SSU rRNA tell us only a fraction of what we need to know, and new surveys, such as Craig Venter’s monumental Global Ocean Sampling expedition (recently reported by Rusch et al., 2007) show that we have consistently underestimated diversity, however this might be defined or measured.

Excitingly, we are starting to have a good idea about how at least some of the apparently rampant oceanic gene swapping occurs: via viruses (Millard et al., 2004; Sullivan et al., 2006). Phages (bacterial viruses), by carrying genes back and forth between alternate hosts, link the evolution of otherwise quite separate lineages. Just how many genes are carried by phages and over what phylogenetic distances is unclear, but there are stunning examples of long (phylogenetic) distance gene transfer in the ocean. Rhodopsin genes have, for instance, been transferred on several occasions back and forth between halophilic archaea and bacteria, and among different phyla of the latter, in aid of light-driven proton or chloride pumping and phototaxis (Sharma et al., 2006).

The more we learn about genetic processes operating among microbes in the ocean, the more confused we get about biological patterns. Extreme, within-species, gene-content variation (some of which clearly is not neutral) uncouples phylogeny (as defined by SSU rRNA) and genetic similarity (as defined by sequence identity of shared genes) from phenotype. Isolates that are very similar genetically and close relatives

phylogenetically can play very different ecological roles. The lack of clear-cut barriers between “species” for information exchange by “lateral” gene transfer or genetic recombination, and the complex interactions between these genetic processes and ecological forces, renders any hunt for a uniform and universally applicable “species definition”—underwritten by any simple or widely relevant “species concept”—quixotic (Doolittle and Papke, 2006) (Figure 5). Thus, there can be no sensible answers to questions, such as how many species are there? How big are their populations? How are they distributed around the globe? We can still study diversity, adaptation, population processes, and microbial biogeography—over a broad range of scales (e.g., Figures 3 and 4)—because the underlying genetic, ecological, and physical dispersal forces are all, in principle, comprehensible and in part comprehended. But we need new concepts about units of biological organization or, better, the courage to develop theory that does not depend on such concepts. Likely, the conceptual focus will be both lower (genes) and higher (communities, consortia, size classes, functional groups) than species. Fortunately, it is this new and dual focus that will be forced upon us (in a good way) by the new science of metagenomics. We will have huge databases linking gene sequences, abundances, and expression levels to oceanic depths, geography, and associated physical, chemical, and biological metadata. We will also have increasingly sophisticated computer models that seek to interpret such data in terms of networks of chemical interactions within and between microbial communities.

What we won't have are "organisms" or "species." The excitement and challenge of metagenomics, which will be the central science of marine microbiology in this new millennium, is to get along without them.

### SIMULATING MICROBIAL COMMUNITY DYNAMICS WITH NUMERICAL MODELS

The revolution in marine microbial genomics and rapid advances in ocean-observation technology will lead to big changes in predictive models of ocean ecosystems (Doney et al., 2004). Already, much progress is being made, but it is not simply a matter of putting more and more information into computer-based simulations of the ocean. Modelers constantly struggle to balance the benefits and drawbacks of complexity versus simplicity in their numerical simulations of marine ecosystems and biogeochemical cycles (Denman, 2003; Hood et al., 2006; Rothstein et al., 2006). Complicated models have many adjustable elements to describe interactions that we know to exist, but predictions are difficult to test because only a few of the model components can be constrained with available observations. Also, computational demands are too high for use in global simulations. Simple models are more manageable and testable, but they cannot effectively address ecological interactions that are not included in their formulations, either explicitly or implicitly. The amazing complexity in marine microbial communities, increasingly revealed by molecular and genomic techniques, thus presents modelers with a formidable challenge (Doney et al., 2004).

A commonly employed strategy,

essentially a compromise between complexity and simplicity, is to model the dynamics of marine organisms as functional groups—groups of organisms or species chosen to reflect size-dependent relationships, biogeochemical functions, and stoichiometric constraints that underlie the patterns observed in nature (Hood et al., 2006; Rothstein et al., 2006). Typically, functional groups distinguish different types of phytoplankton (e.g., very small cyanobacteria, silica-requiring diatoms, carbonate-precipitating coccolithophores, and  $N_2$ -fixers) and large versus small zooplankton. Microbes other than phytoplankton are generally treated as anonymous decomposers of detritus or dissolved organic matter, regenerating nutrients but having no population dynamics of their own. Warranted or not, this simplification is understandable; until recently, marine microbiologists have not been able to assign an ecologically or biogeochemically useful identity to the vast majority of cells (bacteria and archaea) because they could not be isolated and grown in culture. In addition, the biodiversity of nonphotosynthetic marine microbes could not be measured effectively, so their biogeographic patterns could not be described, and the significance of microbial diversity was largely unexplained.

Through the application of new analytical techniques, observations will soon be available to describe the distributions of a broad range of microbes (or at least their genes) whose activities extend well beyond oxygen-producing photosynthesis and aerobic decomposition of organic matter (DeLong and Karl, 2005). If, and only if, the obser-

vations are made with corresponding oceanographic observations and analysis (e.g., Bouman et al., 2006; DeLong et al., 2006; Johnson et al., 2006), the stage will be set to describe the interactions of these organisms with each other and with the physical and chemical environment. But the descriptions will be necessarily complex, and it will be difficult to determine, a priori, how much complexity is needed to describe key drivers of the structure and functioning of ecosystems for the scales being studied and the questions being asked.

Recently, Follows et al. (2007) described a new approach that allows the model itself to determine the degree of complexity needed to describe ecological and biogeochemical interactions in the sea. They employ the equivalent of natural selection "in silico" (by computer simulation) to generate self-organizing, sustainable ecosystems in which the community structure of phytoplankton and the distributions of nutrients and light in the global ocean are determined by the outcome of competition among model organisms randomly chosen from a broad range of physiological types. In their initial study, Follows et al. (2007) simulated phytoplankton by selecting from a palette of capabilities encompassing commonly modeled functional groups and physiologically distinct ecotypes of the numerically dominant marine cyanobacterium *Prochlorococcus*. The results compare well with global patterns of nutrients, estimated primary productivity, and measured distributions of *Prochlorococcus* ecotypes (compare with Johnson et al., 2006); biogeographical distributions can be described with reference to a relatively small number

of physiologically similar groups. The number of groups and their corresponding ecological niches are strongly influenced by the number of environmental variables included in the model, but appropriate complexity in the biota is not imposed, it emerges. This new modeling approach allows us to examine how much detail we must include to address a broad range of questions in marine microbial ecology and biogeochemistry.

In turn, the modeling framework of Follows and colleagues should be useful for assessing the possible ecological roles of uncultured microbes inferred from genomic surveys. The approach would be to assemble model microbes expressing genes detected in nature, allow them to interact in the physical and chemical environment of the virtual ocean, and observe what kind of an ecosystem emerges. Successful simulation of nature—both the survival of microbial types and the evolution of the chemical environment—would depend on many factors, including the development of appropriate rules for constraining the combinations of genetically determined capabilities that could exist in any one microbe. These rules should reflect the stoichiometric requirements, assumed ecological tradeoffs, and constraints on gene transfer that structure microbial communities on the microscopic level and form the basis for the patterns of chemical composition, size structure, and genetic diversity that we observe on larger scales.

### CONCLUDING REMARKS

The intimate relationship between the elemental composition of marine microbes and the chemistry of the

ocean, and the robust consistency in the size structure of marine organisms from microbes to whales illustrate that fundamental patterns in ocean ecology and chemistry are largely emergent from processes operating at lower levels of organization. But until recently, there was no basis for describing in a mechanistic way the ecological and biogeochemical links between processes on the molecular scale all the way to global change. New revelations about microbes in the ocean, particularly descriptions of their incredible genetic diversity, are eliminating the barriers to a comprehensive approach to microbial oceanography. But the revolution has only begun. Soon, physical, chemical, biological, and genomic variability will be observable simultaneously on oceanographic scales that matter. Through interactions across disciplines, observed patterns can be scaled up through statistical analysis and modeling to describe the structure and functioning of marine ecosystems. Interdisciplinary efforts to understand the “sea of microbes” will lead to vastly improved predictions of the ocean’s role in a changing world.

### ACKNOWLEDGEMENTS

We thank the editors and the Gordon and Betty Moore Foundation for giving us the opportunity to work together on this contribution, and Dave Karl for perceptive comments. Charles Stock and Thomas Powell provided material and comments on size spectra. John Cullen was supported by the Natural Sciences and Engineering Research Council of Canada (Research Partnerships, Discovery Grants), the National Ocean Partnership Program, the Office of

Naval Research, and the Killam Trusts. W. Ford Doolittle was supported by the Canada Research Chair Program and the Canadian Institute for Advanced Research. Simon Levin was supported by DARPA, UNESCO, and the Mellon Foundation. William Li was supported by the Department of Fisheries and Oceans, Canada. © Her Majesty in Right of Canada as represented by the Minister of Fisheries and Oceans. 

### REFERENCES

- Arrigo, K.R. 2005. Marine microorganisms and global nutrient cycles. *Nature* 437:349–355.
- Babin, M., J.J. Cullen, C.S. Roesler, P.L. Donaghay, G.J. Doucette, M. Kahru, M.R. Lewis, C.A. Scholin, M.E. Sieracki, and H.M. Sosik. 2005. New approaches and technologies for observing harmful algal blooms. *Oceanography* 18(2):210–227.
- Behrenfeld, M.J., R.T. O’Malley, D.A. Siegel, C.R. McClain, J.L. Sarmiento, G.C. Feldman, A.J. Milligan, P.G. Falkowski, R.M. Letelier, and E.S. Boss. 2006. Climate-driven trends in contemporary ocean productivity. *Nature* 444:752–755.
- Bouman, H., O. Ulloa, D. Scanlan, K. Zwirgmaier, W. Li, T. Platt, V. Stuart, R. Barlow, O. Leth, L. Clementson, and others. 2006. Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* 312:918–921.
- Coleman, M.L., M.B. Sullivan, A.C. Martiny, C. Steglich, K. Barry, E.F. DeLong, and S.W. Chisholm. 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* 311:1,768–1,770.
- Countway, P.D., and D.A. Caron. 2006. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Applied and Environmental Microbiology* 72:2,496–2,506.
- Cullen, J.J. In press. Observation and prediction of harmful algal blooms. In: *Real-Time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling*. M. Babin, C.S. Roesler, and J.J. Cullen, eds, UNESCO, Paris, France.
- DeLong, E.F., and D.M. Karl. 2005. Genomic perspectives in microbial oceanography. *Nature* 43:336–342.
- DeLong, E.F., C.M. Preston, T. Mincer, V. Rich, S.J. Hallam, N.-U. Frigaard, A. Martinez, M.B. Sullivan, R. Edwards, B.R. Brito, and others. 2006. Community genomics among stratified microbial assemblages in the ocean’s interior. *Science* 311:496–503.
- Denman, K.L. 2003. Modelling planktonic ecosystems: Parameterizing complexity. *Progress in*

- Oceanography* 57:429–452.
- Dickey, T.D., M.R. Lewis, and G. Chang. 2006. Optical oceanography: Recent advances and future directions using global remote sensing and in situ observations. *Reviews of Geophysics* 44: RG1001, doi:10.1029/2003RG000148.
- Doney, S.C., M.R. Abbott, J.J. Cullen, D.M. Karl, and L. Rothstein. 2004. From genes to ecosystems: The ocean's new frontier. *Frontiers in Ecology and the Environment* 2:457–466.
- Doolittle, W.F., and R.T. Papke. 2006. Genomics and the bacterial species problem. *Genome Biology* 7:116.
- Droop, M.R. 1974. The nutrient status of algal cells in continuous culture. *Journal of the Marine Biological Association of the United Kingdom* 54:825–855.
- Dubelaar, G.B.J., K. Tangen, P.L. Gerritzen, A.E.R. Beeker, and R.R. Jonker. 1998. CytoBuoy: In situ optical scanning of individual particles with a buoy mounted flow cytometer. Pp. 1,500–1,511 in *Third European Marine Science and Technology Conference, Lisbon, 23–27 May 1998, Project Synopses*. K.-G. Barthel, H. Barth, M. Bohle-Carbonell, C. Fragakis, E. Lipiatou, P. Martin, G. Ollier, and M. Weydert, eds, EUR 18220 EN.
- Eppley, R.W., and B.J. Peterson. 1979. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282:677–680.
- Falkowski, P.G., R. Barber, and V. Smetacek. 1998. Biogeochemical controls and feedbacks on ocean primary production. *Science* 281:200–206.
- Fennel, K., M.J. Follows, and P. Falkowski. 2005. The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic ocean. *American Journal of Science* 305:526–545.
- Follows, M.J., S. Dutkiewicz, S. Grant, and S.W. Chisholm. 2007. Emergent biogeography of microbial communities in a model ocean. *Science* 315:1,843–1,846.
- Fraser-Liggett, C.M. 2005. Insights on biology and evolution from microbial genome sequencing. *Genome Research* 15:1,603–1,610.
- GESAMP. 2001. *A Sea of Troubles*. Reports and Studies. Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP). UNEP, No. 70, 35 pp. Available online at: <http://unesdoc.unesco.org/images/0012/001229/122986e.pdf>.
- Hood, R.R., E.A. Laws, R.A. Armstrong, N.R. Bates, C.W. Brown, C.A. Carlson, F. Chai, S.C. Doney, P.G. Falkowski, R.A. Feely, and others. 2006. Pelagic functional group modeling: Progress, challenges and prospects. *Deep-Sea Research Part II* 53:459–512.
- Hutchinson, G.E. 1961. The paradox of the plankton. *American Naturalist* 95:137–147.
- Johnson, Z.I., E.R. Zinser, A. Coe, N.P. McNulty, E.M.S. Woodward, and S.W. Chisholm. 2006. Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311:1,737–1,740.
- Karl, D.M. 2002. Nutrient dynamics in the deep blue sea. *Trends in Microbiology* 10:410–418.
- Kjørboe, T. 1993. Turbulence, phytoplankton cell size, and the structure of pelagic food webs. *Advances in Marine Biology* 29:1–72.
- Klausmeier, C.A., E. Litchman, T. Daufresne, and S.A. Levin. 2004a. Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature* 429: 171–174.
- Klausmeier, C.A., E. Litchman, and S.A. Levin. 2004b. Phytoplankton growth and stoichiometry under multiple nutrient limitation. *Limnology and Oceanography* 49(4, part 2):1,463–1,470.
- Li, W.K.W. 2007. Macroscopic patterns in marine plankton. In: *Encyclopedia of Biodiversity*. S.A. Levin, ed., Elsevier.
- Li, W.K.W., W.G. Harrison, and E.J.H. Head. 2006. Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proceedings of the Royal Society B* 273:1,953–1,960.
- Longhurst, A.R. 2007. *Ecological Geography of the Sea*, 2<sup>nd</sup> ed. Academic Press, 560 pp.
- Margalef, R., M. Estrada, and D. Blasco. 1979. Functional morphology of organisms involved in red tides, as adapted to decaying turbulence. Pp. 89–94 in *Toxic Dinoflagellate Blooms*. D.L. Taylor and H.H. Seliger, eds, Elsevier-North Holland.
- Martiny, A.C., M.L. Coleman, and S.W. Chisholm. 2006. Phosphate acquisition genes in *Prochlorococcus* ecotypes: Evidence for genome-wide adaptation. *Proceedings of the National Academy of Sciences of the United States of America* 103:12,552–12,557.
- Maynard Smith, J., E.J. Feil, and N.H. Smith. 2000. Population structure and evolutionary dynamics of pathogenic bacteria. *BioEssays* 22:1,115–1,122.
- Millard, A., M.R.J. Clokie, D.A. Shub, and N.H. Mann. 2004. Genetic organization of the psbAD region in phages infecting marine *Synechococcus* strains. *Proceedings of the National Academy of Sciences of the United States of America* 101:11,007–11,012.
- Miller-Rushing, A.J., R.B. Primack, D. Primack, and S. Mukunda. 2006. Photographs and herbarium specimens as tools to document phenological changes in response to global warming. *American Journal of Botany* 93:1,667–1,674.
- Myers, R.A., J.K. Baum, T.D. Shepherd, S.P. Powers, and C.H. Peterson. 2007. Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* 315:1,846–1,850.
- Not, F., M. Latasa, D. Marie, T. Cariou, D. Vaulot, and N. Simon. 2004. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Applied and Environmental Microbiology* 70:4,064–4,072.
- Olson, R.J., A. Shalapyonok, and H.M. Sosik. 2003. An automated subsmersible flow cytometer for analyzing pico- and nanophytoplankton: FlowCytobot. *Deep-Sea Research Part I* 50:301–315.
- Pace, N.R. 1997. A molecular view of microbial diversity and the biosphere. *Science* 276:734–740.
- Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421:37–42.
- Pauly, D., V. Christensen, J. Dalsgaard, R. Froese, and F.C. Torres Jr. 1998. Fishing down marine food webs. *Science* 279:860–863.
- Platt, T., and K. Denman. 1978. The structure of pelagic marine ecosystems. *Journal du Conseil, Conseil International pour l'Exploration de la Mer* 173:60–65.
- Raitsos, D.E., P.C. Reid, S. J. Lavender, M. Edwards, and A.J. Richardson. 2005. Extending the SeaWiFS chlorophyll data set back 50 years in the northeast Atlantic. *Geophysical Research Letters* 32:LO6603.
- Redfield, A.C. 1958. The biological control of chemical factors in the environment. *American Scientist* 46:205–222.
- Rodríguez, J. 1994. Some comments on the size-based structural analysis of the pelagic ecosystem. *Scientia Marina* 58:1–10.
- Rothstein, L., J.J. Cullen, M. Abbott, E.P. Chassignet, K. Denman, S.C. Doney, H. Ducklow, K. Fennel, M. Follows, D. Haidvogel, and others. 2006. Modeling ocean ecosystems: The PARADIGM program. *Oceanography* 19(1):22–51.
- Rusch, D.B., A.L. Halpern, G. Sutton, K.B. Heidelberg, S. Williamson, S. Yooshep, D. Wu, J.A. Eisen, J.M. Hoffman, K. Remington, and others. 2007. The *Sorcerer II* Global Ocean Sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biology* 5:e77.
- Sharma, A.K., J.L. Spudich, and W.F. Doolittle. 2006. Microbial rhodopsins: Functional versatility and genetic mobility. *Trends in Microbiology* 14:463–469.
- Sheldon, R.W., A. Prakash, and W.H. Sutcliffe. 1972. The size distribution of particles in the ocean. *Limnology and Oceanography* 17:327–340.
- Smetacek, V., and S. Nicol. 2005. Polar ocean ecosystems in a changing world. *Nature* 437:362–368.
- Sullivan, M.B., D. Lindell, J.A. Lee, L.R. Thompson, J.P. Bielawski, and S.W. Chisholm. 2006. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biology* 4:e234.
- Sverdrup, H.U. 1955. The place of physical oceanography in oceanographic research. *Journal of Marine Research* 14:287–294.
- Thompson, J.R., S. Pacocha, C. Pharino, V. Klepac-Ceraj, D.E. Hunt, J. Benoit, R. Sarma-Rupavtarm, D.L. Distel, and M.F. Polz. 2005. Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307:1,311–1,313.
- Williams, P.J. LeB. 2006. An appreciation of Alfred C. Redfield and his scientific work. *Limnology and Oceanography Bulletin* 15:53–70.
- Zhu, F., R. Massana, F. Not, D. Marie, and D. Vaulot. 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology* 52:79–92.